1986

Genomic Organization and Sequence Analysis of Rabbit Muscle Phosphofructokinase Gene.

Chan Ping Lee
Louisiana State University and Agricultural & Mechanical College

INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.

- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.

- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17"x 23" black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6"x 9" black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.

Lee, Chan Ping

GENOMIC ORGANIZATION AND SEQUENCE ANALYSIS OF RABBIT MUSCLE PHOSPHOFRUCTOKINASE GENE

The Louisiana State University and Agricultural and Mechanical Col. PH.D. 1986

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages ✓
2. Colored illustrations, paper or print _____
3. Photographs with dark background ✓
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received ✓
16. Other _____

University
Microfilms
International

GENOMIC ORGANIZATION AND SEQUENCE ANALYSIS OF RABBIT MUSCLE PHOSPHOFRUCTOKINASE GENE

A Dissertation
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biochemistry

by
Chan Ping Lee
B.S., National Taiwan University, 1978
December 1986

ACKNOWLEDGMENTS

I would like to express sincere thanks to my major professor Dr. Simon H. Chang for his constant support and guidance throughout the course of my research and the preparation of this dissertation.

I would also like to thank Dr. Hugh D. Braymer, Dr. Jesse M. Jaynes, Dr. Ding S. Shih, and Dr. Ezzat S. Younathan for their valuable suggestions and expert assistance during this study.

Since this project is a joint work, I deeply thank Ming-Ching Kao for his cooperation in this project. I also give thanks to Ben Valdez, Sam Landry, Dr. Susan Payne, Brent French, Michelle Shirley, Michael LoMonaco for their assistance and discussion.

I would also like to deeply thank my family, who never stopped encouraging and supporting me throughout this research.

FORWARD

In order to make this document a useful guide for future investigator in this area, detailed experimental recipes and protocols were described in APPENDICES. Three experimental strategies and techniques developed through this project were included in this dissertation.

TABLE OF CONTENTS

LIST of TABLES

LIST of FIGURES

ABSTRACT

A rabbit genomic clone containing the muscle phosphofructokinase (PFK) gene of approximately 17 kilobase pairs encoding 779 amino acids was isolated and sequenced. This gene contains 22 exons, ranging from 45 to 190 base pairs, and is split by 21 introns of 73-3500 base pairs. Eighty six percent of this gene is occupied by introns. An additional incomplete intron is found in the 5' flanking region, while an Alu-like sequence is identified in the 3' flanking region. The amino acid sequence translated from the coding sequence of this gene reveals 29 residues (amino acid numbers 479-507) which complete a previously unidentified gap [Poorman, R. A. et al., Nature, 309, 467 (1984)] and 4 positions of discrepancy; Ser -> Arg (268), Leu -> Pro (442), Ile -> Ser (558), and the insertion of an additional Arg at position 565. Twelve exons encoding the N-half of the protein are scattered over 13 kbp and the other ten exons encoding the C-half are clustered in a length of 4 kbp. Most introns in this gene occupy positions between or at the ends of the secondary structural elements, but they are not located at identical positions in the two protein-coding halves of the gene. When these exons are identified with their encoded functional subdomains, the exon arrangement shows a duplication pattern between the two halves of the gene. These data support, at the gene level, the hypothesis that mammalian PFK evolved from a prokaryotic progenitor by gene duplication and divergence.

INTRODUCTION

Phosphofructokinase plays a key regulatory role in glycolysis

The glycolytic pathway is an essential metabolic pathway in all living organisms. This pathway degrades glucose to generate energy and also provides building blocks for other synthetic reactions. At least ten enzymes have been found to be involved in this metabolic pathway (Figure 1). Among these enzymes, hexokinase, phosphofructokinase, and pyruvate kinase catalyze the irreversible reactions. Thus these three enzymes would be expected to have regulatory and catalytic roles in this pathway. In general, the enzyme which catalyzes the first irreversible reaction and this reaction is unique to this pathway would be considered as the key regulatory enzyme. Although hexokinase catalyze the first irreversible reaction, glucose to glucose 6-phosphate, its product (glucose 6-phosphate) is not only a glycolytic intermediate. It can also be oxidized by the pentose phosphate pathway to generate NADPH or be converted into glycogen by gluconeogenesis. On the other hand, phosphofructokinase (ATP:ß-D-fructofuranose 6-phosphate 1-phosphotransferase, EC 2.7.1.11(PFK)), which catalyzes the transfer of the gamma-phosphate from ATP to the carbon-1 hydroxyl of D-fructose 6-phosphate to form ADP and D-fructose 1,6-biphosphate, is the first irreversible reaction unique to the glycolytic pathway. Thus, PFK is highly appropriate to be the key regulatory enzyme in glycolysis (Passonneau & Lowry, 1962).

Fructose-bisphosphatase (FBPase) is the antagonist of PFK

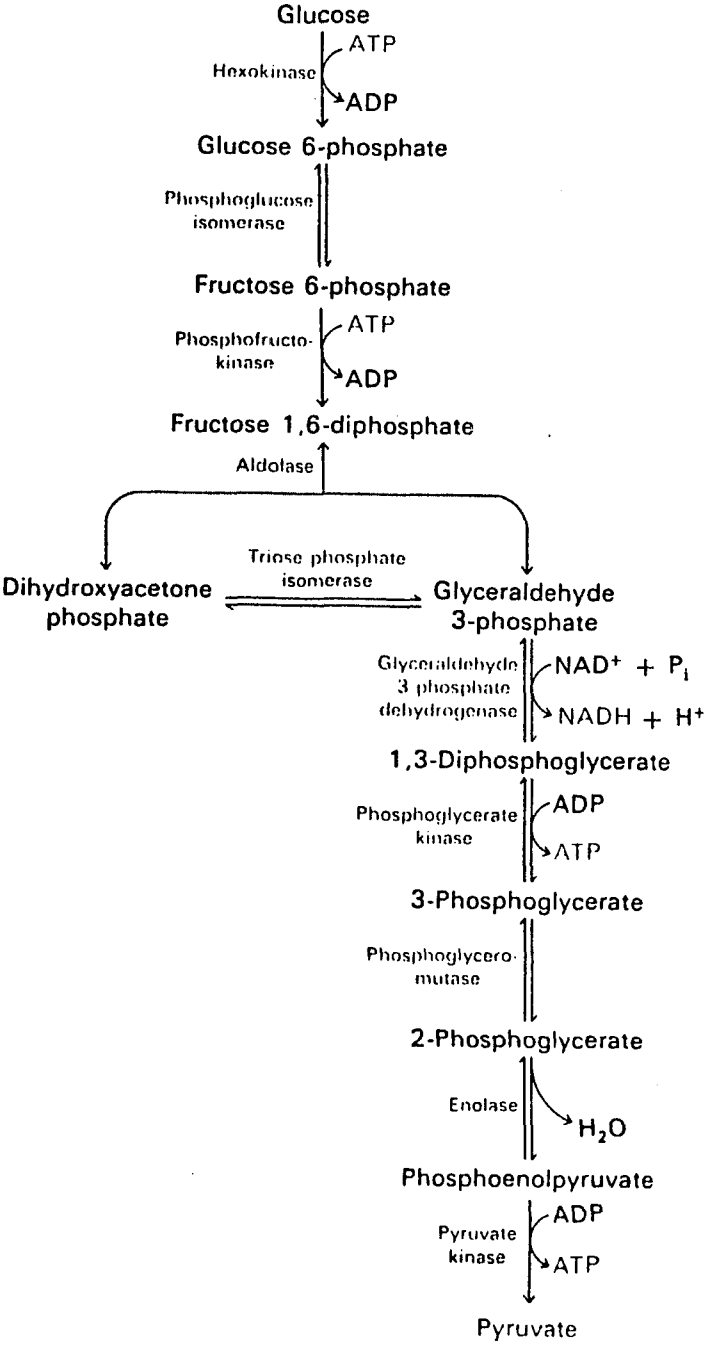

Figure 1. The glycolytic pathway (taken from Stryer, 1981)

Frucotse-bisphosphatase catalyzes the decomposition of fructose 1,6-bisphosphate into fructose 6-phosphate by hydrolysis of the phosphate ester at C-1 in gluconeogenesis. This is an opposite irreversible reaction of PFK. It seems that PFK and FBPase perform a futile cycle; phosphorylation of fructose 6-phosphate to fructose 1,6-diphosphate and hydrolysis back to fructose 6-phosphate. However, both reactions are not simultaneously fully active in most cells because of reciprocal allosteric controls. Koerner et al. (Koerner et al., 1977) proposed a regulatory mechanism for both enzymes. In this mechanism, PFK converts ß-D-fructose 6-phosphate (beta-F6P) to ß-D-fructose 1,6-bisphosphate (beta-FBP). This reaction is activated by alpha-D-fructose 1,6-bisphosphate (alpha-FBP; the anomer of beta-FBP), but is inhibited by alpha-D-fructose 6-phosphate (alpha-F6P; the anomer of beta-F6P). In contrast, FBPase, which is activated by beta-F6P but is inhibited by beta-FBP. converts alpha-FBP to alpha-F6P. This regulatory mechanism thus leads to reduction in the rate of futile cycling. Similar to FBPase, PFK is subject to allosteric activation and inhibition by various metabolites. For example, ATP and citrate are the inhibitors and ADP and AMP are the activators for mammalian PFK. When the cell needs both energy (ATP) and building block (citrate), PFK is in the active condition. When the cells are abundant in ATP and citrate, PFK would be inhibited by these two effectors.

Mammalian phosphofructokinases are tissue-specific isozymes

Multiple forms of PFK from various mammalian sources have been reported. At least four forms of PFK have been found in normal and tumor tissues of human (Layzer and Conway, 1970) and rats (Tanaka et

al., 1971). Human PFK isozymes are under the control of three structural loci (Kahn et al., 1979). It has been shown that there are three isozymes in rabbit (Tsai and Kemp, 1973): type A found in skeletal muscle and heart; type B found in erythrocytes and liver; and type C found along with A and B in brain, thymus, and several other tissues. In addition, lung, adipose tissue, and stomach have an A-B hybrid set. Each isozyme is composed of four identical subunits. The reported molecular weights of these subunits are 84,000 for A subunit, 80,000 for B subunit, and 86,000 for C subunit (Foe and Kemp, 1984).

One-dimensional chymotryptic and staphylococcal V8 pretease fingerprinting analyses show that the three isozymes have different structure (Foe and Kemp, 1986). All three isozymes exhibit a sigmoidal response to increasing F6P concentration and sensitivity to inhibition by increasing concentration of ATP but they are kinetically distinct one another. Isozyme C is more sensitive than isozyme A but less sensitive than the isozyme B to inhibition by ATP, and less sensitive than isozyme A but more sensitive than isozyme B to inhibition by citrate (Foe and Kemp, 1986). These isozymes have different self-association property. When they are centrifuged on sucrose gradients, isozyme A and B will self-associate to oligomers higher than tetramer, but isozyme C will not (Foe and Kemp, 1986). Immunoprecipitation experiments indicate that these three isozymes have different antigenic determinants. Isozyme C will not precipitate with the antiserum raised against isozyme A or isozyme B.

Random tetremerization of these subunits may produce various homo- and heterotetrameric isozymes. These hybrid isozymes are also distinguishable from one another by the difference in their self-association,

one dimensional peptide mapping, immunoprecipitation and regulatory kinetic properties (Foe and Kemp, 1986). Thus the separate properties of isozymes A, B, and C cannot be used to predict the properties of hybrid isozymes.

The amino acid sequence of rabbit muscle PFK

Rabbit muscle PFK (RMPFK), isozyme A, has been studied for a long time since this enzyme revealed attractive allosteric features for enzymologists (Hofmann, 1978). A nearly complete amino acid sequence of 749 residues has been reported (Poorman et al., 1984). However, the reported amino acid sequence had an undetermined gap of about 30 residues in the C-terminal region. One of the features for this undetermined gap is that the correspondant peptide is very hydrophobic (Kemp, unpublished data). This feature makes the purification of this peptide difficult. More importantly, this gap comprises part of the substrate ATP binding site in the C-terminal of RMPFK. Identification of residues in this gap is important for understanding the function of the C-terminal of RMPFK.

The features of Bacillus stearothermophilus PFK

Among PFKs, Bacillus stearothermophilus PFK(BsPFK) is the most well studied case. BsPFK is a tetramer of identical subunits each of molecular weight 36,000. The complete amino acid sequence of 316 residues has been determined (Kolb et al., 1980). Kinetic analysis reveals that BsPFK has a sigmoidal response in the presnece of substrate F6P and is allosterically regulated by ADP, an activator, and phosphoenol pyruvate (PEP), an inhibitor. As compared to RMPFK, BsPFK is about

half the size of RMPFK and is regulated by a less effector. BsPFK is not inhibited by ATP and citrate, nor activated by AMP and fructose bisphosphate (F6P). The crystal structure of the active conformation has been solved to 2.4 Å resolution (Evans and Hudson, 1979) and inactive conformation has also been sloved to 7 Å resolution (Evans et al., 1986). Three ligand-binding sites (ATP, F6P, and ADP binding sites) have been located in this crystal structure (Figure 2). The active site includes substrate ATP and substrate F6P binding and the effector site includes activator ADP and inhibitor PEP binding. These are three distinct features of these binding sites: (1) the position of gamma-phosphate of ATP is very close to the 1-hydroxyl of F6P, (2) the binding of the phosphate of F6P involves two arginines from a neighboring subunit in the tetramer; and (3) the activator ADP is also bound by residues from two subunits. The first feature reveals that there is a suitable position for in-line gamma-phosphate transfer from ATP to F6P. The second and third feature suggest that there are rearrangements of the subunits when the enzyme reacts with substrates or effectors and these rearrangements will result in the conformational change of enzyme (Evans et al., 1981).

The features of the hypothetic RMPFK protein model

Strong homology has been observed between BsPFK and rabbit muscle PFK and between the N- and C-halves of rabbit muscle PFK (Poorman et al., 1984; Figure 3). The respective identities for each comparison are 44% (BsPFK vs N-half), 34% (BsPFK vs C-half) and 32% (N-half vs C-half). Based on comparative sequence analysis, X-ray crystallography of BsPFK, chemical modification studies, kinetic analysis and limited

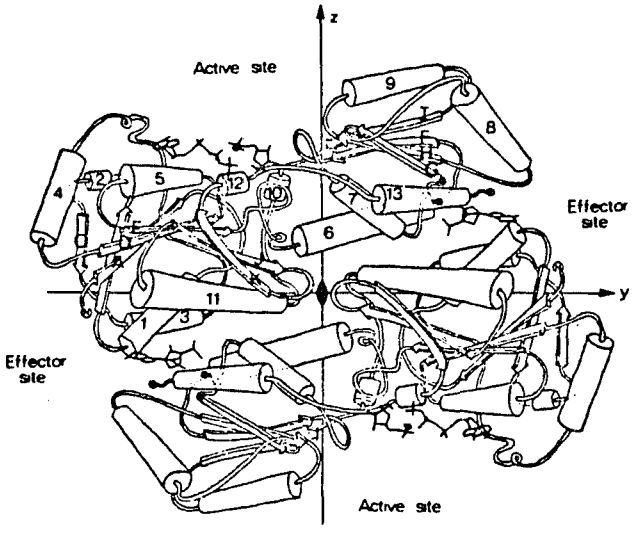

Figure 2. **Schematic views fo two subunits along the x-axis in the BsPFK tetramer.** ß-strands are indicated by arrows (A-K) and alpha-helices by cylinders (1-13). The subunits ATP and F6P are shown in the active sites and the activator ADP in the effector sites (taken from Evans & Hudson, 1979).

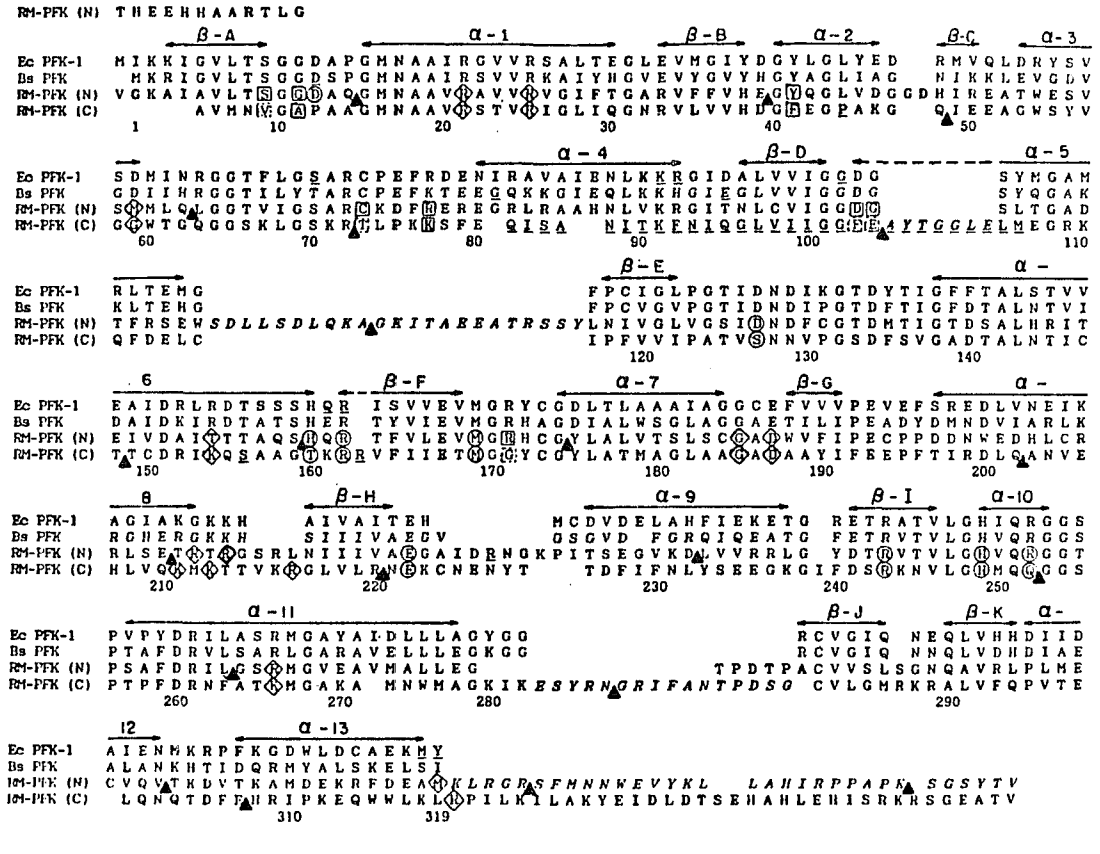

Figure 3. Homology comparison of related PFKs and position of introns along the amino acid sequence of rabbit muscle PFK. The amino acid sequence of E. coli PFK (Ec PFK-1), B. stearothermophilus PFK (Bs PFK) and the N-half (RM-PFK (N) and C-half (RM-PFK (C)) of rabbit muscle PFK are aligned for maximum homology. Double headed arrows above the sequences indicate secondary structural elements found in B. stearothermophilus PFK. Introns (solid triangles) in the N- and C- halves of rabbit muscle PFK are positioned between the respective amino acid residues. Underlined regions are corrected or newly identified amino acid residues in all three PFK sequences. Residues boxed with square boxes are binding sites for substrate ATP; diamond-shaped boxes, binding sites for ATP inhibitor or ADP activator, circle boxes are binding sites of F6P and FBP (modified from Hellinga and Evans, 1985 with permission).

proteolysis data, Poorman and his coworkers proposed a protein model for RMPFK (Poorman et al., 1984; Figure 4). The major assumptions in this model are:

(1) The evolution of proteins proceeds with conservation of three-dimentional structure, despite wide-ranging diversification of amino acid sequence. Similar to the subunit structure of BsPFK, each half of RMPFK would also consist of two domains (Figure 2). Each domain has a central ß-sheets core sandwiched by alpha-helices (alpha/beta class protein (Levitt and Chothia, 1976)). The first domain has seven strands of ß-sheet, the central five being parallel, and the outer two antiparallel to the others. Alpha-helices 1 to 5 and 10 to 12 surround the central ß-sheets core. The second domain has four parallel strands. Alpha-helices 6 to 9 and 13 surround the central ß-sheet core. The two ß-sheets cores point towards a deep cleft which forms the active site. The substrate ATP is bound almost entirely by domain 1 of N-half. This site lies between helix 5 (residue 102-108, N-half; Figure 3) and a surface loop between helices 3 and 4 (residues 72-77, N-half). The F6P molecule is bound almost entirely by domain 2 of N-half. The most distinct feature for this site is the involvement of residues from the neighboring subunit. The 6-phosphate group is bound by His-249 and Arg-252 from its own subunit, and by Arg-162 and Arg-243 from the neighboring subunit (Figure 4 and Table 1).

(2) The N- and C-halves of RMPFK monomer are joined by a connecting peptide of about 30 residues (Lys 320, N-half to Ala 5, C-half; Figure 3). This connecting peptide brings the N- and C-halves together and reslts in the interaction of these two halves (Figure 4).

(3) A RMPFK tetramer is roughly correspondant a BsPFK octomer and

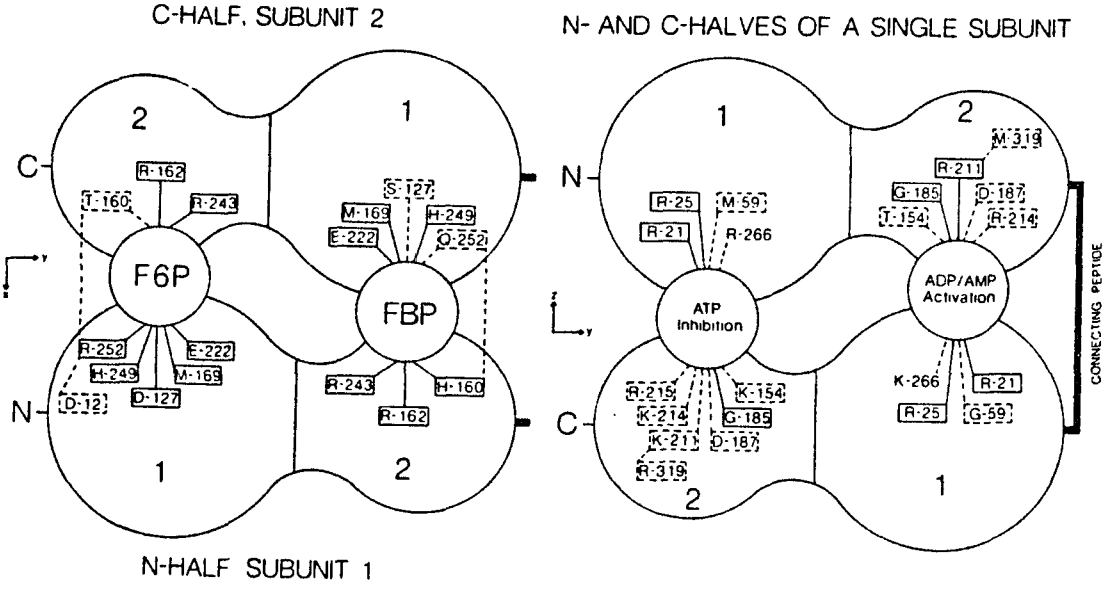

Figure 4. **Location of substrate and effector sites in the hypothetic RMPFK protein model.** Residues enclosed in solid boxes are identical to those at corresponding positions in BsPFK; those enclosed in dotted boxes are different. The function of each residue is explained in Table 1. (taken from Poorman et al., 1984).

Table 1. Comparison of residues in the ligand binding sites for BsPFK and N- and C-halves of RMPFK. (modified from Hellinga and Evans, 1985)

BsPFK	Function	RMPFK(N)	RMPFK(C)
substrate F6P binding		substrate F6P binding	activator FBP binding
R-162	binds 6-P	R	R
R-243	binds 6-P	R	R
H-160	subunit interface	H	T
D-127	catalytic residue	D	S
M-169	hydrophobic contact to sugar ring	M	M
E-222	H-bond to O4	E	E
H-249	binds 6-P	H	H
R-252	H-bond to P6, O2, and O5 also central to subunit interface	R	Q
N-12	subunit interface	D	P
T-156	subunit interface	T	I
substrate ATP binding		substrate ATP binding	-----
S-9	contact to ribose	S	V
G-11	main-chain amide H-bond to gamma-P	G	A
Y-41	contact to ribose	Y	F
C-73	main-chain amide and carbonyl H-bond to ribose hydroxyls	C	T
K-77	hydrophobic contact to adenine	R	K
D-103	possible interaction with Mg	D	F
G-104	main-chain amide H-bond to ß-P, carbonyl to O1'. No room for side-chain	G	E
Q-107	possible H-bond to adenine N6	T	E
K-111	hydrophobic contact to adenine	T	Q
R-71	catalytic residue (binds gamma-P)	R	G
activator ADP binding		activator ADP binding	inhibitor ATP binding
R-21	H-bond to ß-P	R	R
R-25	H-bond to alpha- and ß-P	R	R
D-59	main-chain amide H-bond to ß-P side-chain H-bond to ribose O3' and Arg-63	M	G
R-63	H-bond to Asp-59, almost the only contact across the y-dyad	L	Q
R-154	H-bond to ß-P	T	K
G-185	main-chain carbonyl H-bond to Mg	G	G
E-187	carbonyl H-bonds to Mg	D	D
R-211	H-bonds to alpha-P and adenine N7, coupled with C terminus	T	K
K-213	H-bonds to alpha-P	T	K
K-214	main-chain amide H-bond to ribose O1'	R	T

would be expected to have eight active sites and eight allosteric sites. However, equilibrium binding studies have shown that there are only four catalytic sites in RMPFK tetramer. Also, RMPFK is activated by FBP and inhibited by ATP but BsPFK is not. These properties lead to the assumption that RMPFK have mutated to generate new effector sites and have lost half of the active sites (Figure 4). Different from BsPFK, the hypothetic protein model reveals an activator FBP binding site and an inhibitor ATP binding site in each subunit. The activator FBP binding site is assumed to have mutated from substrate F6P binding site by replacement of the catalytically essential Asp-127 with a serine residue (Figures 3 and 4). Asp-127 functioned as a base catalyst to increase the nucleophilicity of the O-1 hydroxyl of F6P for attack on the gamma-phosphate of ATP (Table 1). The replacement of Asp-127 by Ser in the new activator FBP binding site would remove this negative change and provide more room for the binding of FBP.

(4) Arg-71 is a catalytic residue to bind gamma-P in substrate ATP binding site. This residue is conserved in N-half and has been mutated to Gly in C-half. The C-half may lose the activity of substrate ATP binding due to this mutation. Asp-103, and Gly-104 are also important in substrate ATP binding, but the corresponding residues in C-half of RMPFK were unknown at that time.

(5) Inhibitor ATP binding site is assumed mutating from activator ADP binding site although the reasons are not clear.

Gene duplication and divergence

RMPFK is a tetramer of identical subunits each of which is about twice the size of those from BsPFK. Both RMPFK and BsPFK are allo-

steric enzymes, but RMPFK has more complex allosteric properties. Allignment of the amino acid sequence of RMPFK with BsPFK shows that RMPFK is very similar to the BsPFK (Poorman et al., 1984; Figure 3). In addition, the N-half and C-half of RMPFK share an internal sequence homology (Figure 3). This evidence leads to the speculation that mammalian PFK evolved from a prokaryotic progenitor by gene duplication and divergence (Poorman et al., 1984).

Rationales

Obviously, the hypothetical RMPFK protein model and speculated pathway of PFK gene evolution provides a good basis for the studies of RMPFK. However, there is still no experimental evidence to prove or disprove this hypothesis. Studies on the structural organization of the rabbit muscle PFK gene would therefore reveal the following significance:

(1) To test the speculated pathway of gene evolution.

(2) To complete the amino acid sequence of RMPFK.

(3) To provide a molecular basis for constructing a full-length cDNA. This full-length cDNA can be subcloned and expressed in E.coli cells and the sequences corresponding to those binding sites can be changed by site-directed mutagenesis. Mutated PFK can be made by this way and used to examine the hypothetical protein model.

(4) Homozygous deficiency of the muscle PFK in man results in glycogen storage disease (Vora et al., 1985), which is characterized by exertional muscle weakness and compensated hemolysis. It has also been found that the level of PFK specific activity of diabetic rats was lower than that from control rats (Bauer & Younathan, 1984). The

causes of these PFK deficiency and abnormality are still unknown. The RMPFK gene may serve as a good probe to approach these clinical problems.

Based on these rationales, we isolated the gene for rabbit muscle PFK and determined its DNA sequence. The genomic organization and sequence features of this gene will be presented in this dissertation.

The following three techniques are develpoed in this work. Their rationales are presented here:

Restriction mapping strategy

A detailed restriction map is very helpful for DNA sequencing and gene characterization. Established procedures for restriction mapping involve either isolation of a large amounts of target DNA and elution of radioactive DNA from agarose gel (Smith & Birnstiel, 1976) or repeat gel electrophoresis and hybridization. One strategy uses synthetic oligonucleotides for selective end labelling (Rackwitz et at., 1984), however this method is only suitable for lambda phage vector. In order to facilitate DNA sequencing and gene characterization, we have developed a simple and direct reading method for restriction mapping. Thirteen restriction enzymes were successfully used to determine a detailed restriction map of PFK gene.

Specific DNA probes

The design of M13 phage vectors (Vieria & Messing, 1985) allows the isolation of single stranded DNA. Distinguishing which carried an insert in the vegetative phage strand usually forms an experimental bottle neck. The ideal probe to select M13 phage clones is the free

complementary strand without vector sequence. Such fragments are normally isolated by digesting the DNA of interest with restriction endonucleases followed by gel electrophoresis under denatured conditions. Resolution of the strands by this method is usually poor. A technique has been developed for making strand and length specific probes. The idea is partly based on concepts published by Dale (Dale et al., 1986). The single stranded inserts can be obtained by digestion with restriction enzymes at specific sites. These specific sites are formed by two synthetic oligonucleotides. One oligonucleotide is complementary to the region of EcoRI site and the other is complementary to the region of HindIII site. This single stranded DNA can be radioactively labelled at 5' or 3' end or homogeneously labelled on its complementary strand. These strand and length specific probes have been used to efficiently screen our M13 library.

DNA sequencing strategy

Currently existing DNA sequencing strategies can be classified into two groups: random and nonrandom sequencing. The Bal-31 method (Poncz et al., 1982; Misra, 1985) is the well accepted one for nonrandom sequencing, while the sonication method (Deininger,1983) represents the other strategy. The sonication approach possesses the following disadvantages: (1) redundancy in sequencing and analysis is time consuming; (2) multiple inserts in one vector may result in interference during sequence overlapping analysis. Unfortunately, the traditional Bal-31 methods also have disadvantages since large amounts of DNA must be purified from gels (Poncz et al., 1982) and low efficiency of blunt end ligation and positive transformation is a problem

(Misra, 1985). In the case of the 17 kilobase pairs PFK gene, DNA sequencing rapidly became the limiting step. In order to sequence this long gene more efficiently and more economically, we improved the existing Bal-31 methods. This improved sequencing strategy has been successfully used to sequence the PFK gene fragments and will be presented in this dissertation.

MATERIALS and METHODS

I. Materials

A. Chemicals

Agarose, ammonium persulfate, ethidium bromide, and low melting point agarose used for gel electrophoresis were purchased from Bethesda Research Laboratories. Acrylamide,N',N'-Methylenebisacrylamide,N,N,N',N'-tetramethylethylenediamine (TEMED), sucrose, and urea used for DNA sequencing were purchased from New England Biolabs. Inc. Amberlite (MB-1) ion exchange resin was purchased from Mallinckrodt. Bacto tryptone, Bacto agar, yeast extract, nutrient broth used for bacterial growth were purchased from Difco. Antibiotics (ampicillin, chloramphenical, tetracycline) were purchased from Sigma. 5-Bromo-4-chloro-3-indolyl-ß-D-galactoside (X-gal) and isopropyl-ß-D-thiogalactopyranoside (IPTG) used for M13 cloning were purchased from Bethesda Research Laboratories. Dithiothreitol, Hexamine Cobolt (III) chloride, Rubidium chloride used for transformation were purchased from Alfa. Dideoxynucleotides and deoxynucleotides used for DNA sequencing were purchased from Pharmacia. The detailed preparation and composition of reagents were described in APPENDIX 1.

B. Radioactive Materials

Adenosine 5'-triphosphate, tetra(triethylammonium) salt,[gamma-^{32}P] was purchased from New England Nuclear (cat# NEG-002 Z) with the specific activity ~500 ci/mmol. Deoxycytidine 5'-triphosphate, [alpha-^{32}P] was purchased from ICN Biomedical Inc. (cat# 3302H) with

the specific activity > 600 ci/mmol. Deoxyadenosine 5'-[alpha-thio] triphosphate,[^{35}S] was purchased from New England Nuclear (cat# NEG-034H) with the specific activity ~10 mci/ml.

C. Enzymes

Deoxyribonuclease I (DNase I), DNA polymerase I (E. coli), nuclease Bal-31, ribonuclease T_1, T_4 polynucleotide kinase, terminal deoxyribonucleotidyl transferase (TdT) were purchased from Bethesda Research Laboratories. Restriction enzymes and T_4 DNA ligase were purchased from Bethesda Research Laboratories or International Biotechnologies, Inc. Lysozyme and ribonuclease A were purchased from Bethesda Research Laboratories. Pronase and proteinase K were purchased from Boehringer Mannheim Biochemicals. The large fragment of DNA polymerase I (Klenow fragment) was purchased from Boehringer Mannheim Biochemicals or Pharmacia or United States Biochemical Corporation. The preparation of enzymes and composition of buffers were described in APPENDICES 2 and 3.

D. Oligonucleotide

All the oligonucleotides used as probes or primers were synthesized using solid support phosphite chemistry (Matteucci & Caruthers, 1981) in an automated DNA synthesizer (Applied Biosystems, Model 380A). The name and sequence for each oligonucleotide is listed in APPENDIX 4.

E. Bacterial Strain and Vector

E. coli K803 was used as the host cell for lambda Charon 4A.

JM83 was used as the host cell of pUC plasmids. HB101 was used as the host cell of pBR plasmids. JM107 was used as the host cell of M13mp10/11 and M13mp18/19. The bacterial genotype and the vector size were listed in APPENDIX 5.

II. Instrumentation and Equipment

All the important instruments and equipment used in this work were listed in APPENDIX 5.

III. Methods

Isolation and identification of rabbit muscle PFK gene

A rabbit genomic library clonded in lambda Charon 4A phage vector was a gift from Dr. R. C. Hardison (Maniatis et al., 1978). An approximately 200 bp long cDNA of rabbit muscle PFK was obtained from Dr. S. D. Putney (Putney, et al., 1983). The cDNA insert was further subcloned into pBR322 plasmid DNA, nick-translated in the presence of (alpha-^{32}P) dATP and used to probe the rabbit genomic library. Positive phage plaques shown to have reproducible patterns on duplicated filters were isolated and further purified by replating (Chang et al., 1984). Lambda Charon 4A/pfk DNA was prepared through cesium chloride gradient as described (Blattner et al., 1977). Purified phage DNA was digested with restriction endonucleases and fractionated by electrophoresis in 0.7% agarose gels. DNA was transferred bidirectionally onto nitrocellulose filters. Blots were then hybridized to ^{32}P-labelled cDNA probe and synthetic oligonucleotide probe as described by Southern (Southern, 1975). Hybridization condition using synthetic oligonucleotide probe and experimental details for purification of oligonucleotide are given

in APPENDIX 7.

Preparation of plasmids and M13 RF DNAs

Plasmids were grown in E. coli strain HB101. M13 clones were propagated in E. coli strain JM107. Plasmid and M13 RF DNAs were prepared using an alkaline lysis procedure (Birnboim & Doly, 1979) and subsequent purification on cesium chloride ethidium bromide gradients (Maniatis et al., 1982). Full experimental details are given in APPENDIX 7.

Restriction endonuclease mapping

M13 RF DNAs containing the desired subcloned fragments were linearized and radioactivelly labelled at 5' end with T_4 DNA kinase or 3' end with T_4 deoxyribonucleotidyl transferase. The labelled DNAs were partially digested with individual endonucleases that recognize 6-bp sequences (EcoRI, SacI, KpnI, SmaI, XmaI, BamHI, XbaI, SalI, HincII, PstI, SphI. HindIII, and PvuII), and the digests were analyzed by electrophoresis on 1.2% agarose gels. Gels were dried and exposed under an X-ray film at room temperature overnight. This autodiogram yielded a restriction map for each subclone. The position and orientation of individual subclones were determined by the results of hybridization. The principle of this mapping strategy are described in RESULT. The full experimental details including 5' and 3' ends labelling and partial digestion are given in APPENDIX 7.

Cloning and transformation

Cohesive-ended DNA fragments were cloned into plasmid pBR325 and

M13mp18 and 19 by forced cloning. Blunt-ended DNA fragments generated by PvuII digestion were cloned by ligation directly into SmaI site of bacterial alkaline phosphatase treated M13mp18 replicative form DNA. Ligated DNAs were transformed into HB101 for plasmid pBR325 and JM107 for M13mp18 and 19. Full expermental details for the preparation of competent cell and transformation are given in APPENDIX 7.

Assay for M13 recombinants

Individual recombinant from the libraries were grown in YT broth, and phage RF DNA was isolated as described (Birnboim & Doly, 1979). The yield of RF DNA was adequate for insert size determination. Aliquotes of single stranded phage from individual cultures were screened for insert by direct gel electrophoresis and for insert polarity by complementary test as described (Messing, 1983). Full experimental details for direct gel electrophoresis and complementary test are given in APPENDIX 7.

Dot-blot assays

Recombinant M13 phage clones were identified using filter hybridization. One ul of aliquot from the supernatant of phage stock solution were spotted in a grid of nitrocellulose paper. Those clones homologous to a particular DNA fragment were identified by DNA.DNA hybridization. Full experimental details for dot-blot hybridization are given in APPENDIX 7.

DNA sequencing

DNA fragments were sequenced as recombinants in M13mp18 or 19 by

the dideoxy-chain-terminator method (Sanger et al., 1977) using single stranded M13 phage DNA as templates and (alpha-^{35}S)dATP as radioactive label (BNiggin et al., 1983). The primers used in these DNA sequence reactions were M13 universal primers or oligonucleotides corresponding or complementary to the coding regions of rabbit muscle PFK gene. Full experimental details from the preparation of template DNA to sequencing reactions are given in APPENDIX 7.

Sequence analysis

DNA sequence data were assembled and analyzed using the Staden programs (Staden, 1980) on a VAX 11-750 computer.

RESULTS

I. RMPFK gene --- gene identification, sequence analysis, and genomic organization.

1. Identification of rabbit muscle PFK gene.

Purified lambda Charon 4A/pfk DNA was digested with restriction endonucleases. The results were shown in Figure 5. EcoRI digest on lane 4 showed 5 bands: 23 kbp, 17 kbp, 8.9 kbp, 6.0 kbp, 2.5 kbp. The first two bands came from vector as compared to the known restriction map of charon 4A vector. The total length of the insert in this clone is therefore estimated to be 17 kbp. The results also showed that the insert contained three EcoRI sites.

Purified phage DNA was further digested respectively with EcoRI, PstI or PvuII, and hybridized with synthetic N-probe and cDNA probe. Results were shown in Figure 6. On panel (A) lane 7, a 0.5 kbp PvuII fragment was hybridized to the synthetic N-probe, while 2 positive bands of PstI digest were shown on lane 5 of Panel (B). These positive fragments were subcloned into M13 and sequenced. Figure 7 shows the hybridization results of the cloning of C-terminal fragment. All the clones were shown to contain the C-terminal fragment as compared with the negative and positive controls. Template DNAs of both strands were then purified and sequenced. DNA sequence of clone A corresponded to nucleotides 15922-16154, DNA sequence of clone B corresponded to nucleotides 16254-16506. The translated amino acid sequence

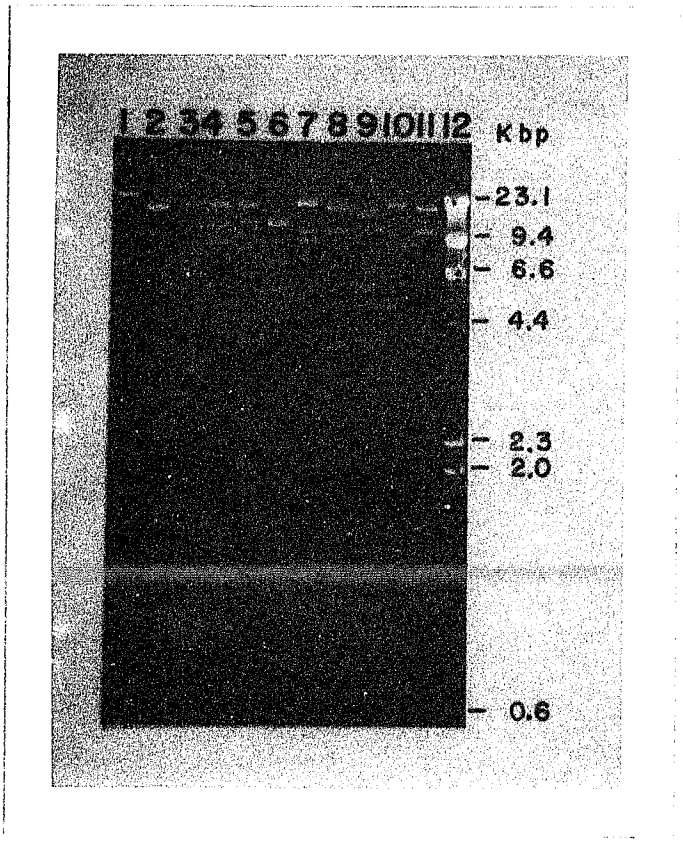

Figure 5. **Restriction endonuclease digestions of lambda Charon 4A/pfk.** Phage DNA, purified from the clone selected by plaque hybridization, were digested with restriction endonucleases. The digests were subjected to electrophoresis on 0.7% agarose gel and stained by ethidium bromide. HindIII digest of wild type lambda DNA was run on Lane 12 as the size markers. The numbers on the right margin indicate the size and position of the marker fragments in kilobase pairs. Lane 1, intact; lane 2, KpnI digest; lane 3, KpnI and EcoRI digests; lane 4, EcoRI digest; lane 5, BamHI digest; lane 6, KpnI and BamHI digests; lane 7, HIndIII digest; lane 8, KpnI and HindIII digests; lane 9, BamHI and EcoRI digests; lane 10, HindIII and EcoRI digests; lane 11, BamHI and HindIII digests.

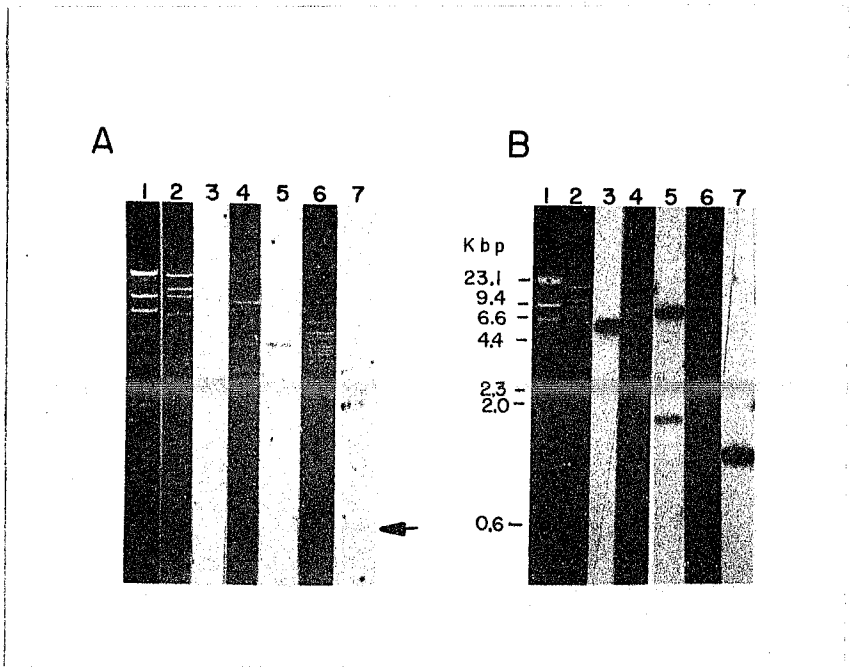

Figure 6. Southern transfer and hybridization to probe the C- and N-terminal of rabbit muscle PFK gene. (A) Probing of N-terminal with synthetic N-probe. Phage DNA from lambda Charon 4A/pfk were digested with EcoRI (lane 2), PstI (lane 4), PvuII (lane 6). HindIII digest of wild type lambda DNA were run on lane 1 as the size markers. 0.7% agarose gel were electrophoresed, blotted, hybridized with the synthetic N-probe as described in MATERIALS and METHODS. The hybridization pattern for each digest was arranged side by side. The arrow indicates the positive band of 0.5 kbp. (B) Probing of C-terminal with cDNA probe. All the arrangements and conditions were described in (A) except hybridization using the nick-translated cDNA probe.

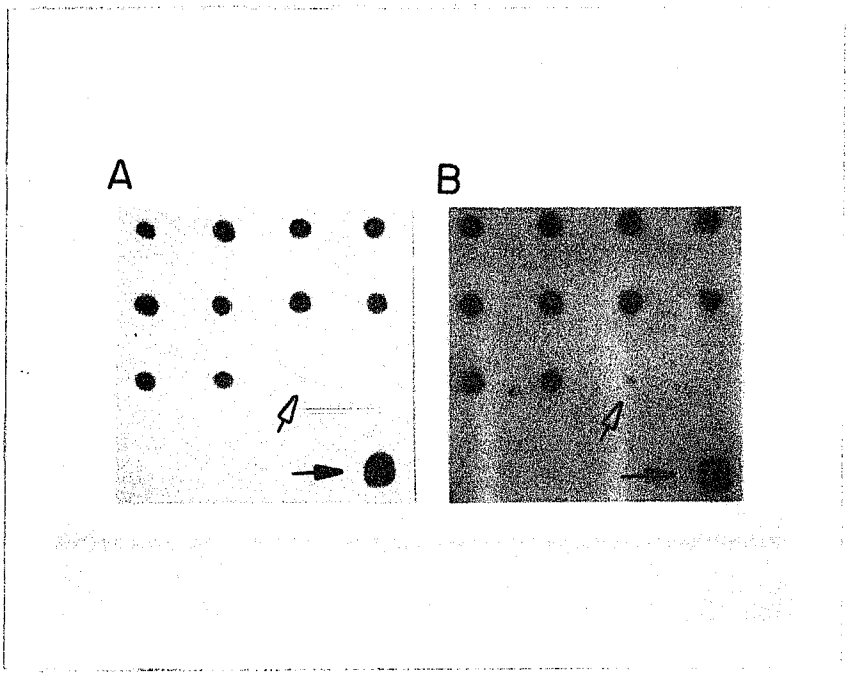

Figure 7. **Dot hybridization to identify M13 clones containing the C-terminal of rabbit muscle PFK gene.** A 0.5 kbp EcoRI/PstI fragment selected from hybridization of lambda Charon 4A/pfk to cDNA probe was subcloned into M13mp18 and 19 respectively. Phage particles were selected for insertion by their lac phenotypic change. 10 putatively positive clones from each vector source were chosen, spot on nitrocellulose paper, and hybridized with the nick-translated 0.5 kbp fragment as described in MATERIALS and METHODS. Blue phage particles and pure insert DNA fragment were spot on the same nitrocellulose paper as the negative and positive controls respectively. (A) M13mp18 was used as vector. (B) M13mp19 was used as vector. The open arrows indicate negative control; whereas the solid arrows show the positive control.

perfectly matched amino acids 725-732 and 766-779 (Figure 8). On the other hand, the 0.5 kbp PvuII fragment hybridized to synthetic N-probe was subcloned into M13. Figure 9(B) showed that two selected clones had slower migration as compared with control. Panel (A) showed that both clones contained the inserts of same size corresponding to 0.5 kbp. Panel (C) showed that these two clones had complementary inserts. Template DNAs purified from these 2 clones were sequenced. The deduced amino acid sequence matched the first 28 amino acids of RMPFK (Figure 8). These data showed that the 17 kbp insert contained not only the C-terminal but also the N-terminal of PFK gene.

2. Restriction endonucleases mapping of PFK gene.

Figure 10 is the schematic representataion of this new strategy. The procedures are represented systematically from step 1 through step 6. A and B represent restriction sites in polylinker. In step 1, the target DNA is subcloned into restriction sites of polylinker of M13 or pUC series vector. In step 2, RF DNA is prepared and then linearized by enzyme B which won't cut the inside of target DNA. In step 3, linearized DNA is radioactivelly labelled at 5' or 3' end. In step 4, the labelled DNA can be digested with enzyme A and then go to step 5 or leave intact for the next step. In step 5 and 5', the labelled DNA is partially digested with different restriction enzymes. In step 6 and 6', mixture is subjected to agarose gel,

Figure 8. The nucleotide sequence of the rabbit muscle PFK gene. The sequence is from the site of initiation of translation (numbered with +1) to the site of termination of translation (numbered with 16300). The 61 bp of 5' flanking region are numbered from -1 to -61 and 583 bp of 3' flanking region are indicated by numbers from 16301 to 16884. The number of base pairs indicated in angle brackets showing the sequences which have not yet been overlapped. Letter X represents uncertained base. The numbers in the right margin correspond to the terminal nucleotide of that sequence. All exons and introns are indicated with a title. The encoded amino acids are numbered above sequence. Splice junctions are indicated by emphasized italic. The restriction sites are indicated by each name under the first base. The putative Lariat branch sites are both numbered and indicated with asterisk. The putative Alu concensus sequence and direct repeats are also designated with arrows below the sequence.

```
     INTRON 0                    26
-----gctagcggcaccttgacccagctgtcctaactgaccatcgtcttgccttctagag/tggatc      -1
    Linker              PvuII    *                  XbaI

EXON 1
    001
Met Thr His Glu Glu His His Ala Ala Arg Thr Leu Gly Val Gly Lys Ala
ATG ACC CAT GAA GAG CAC CAT GCA GCC AGA ACC CTG GGG GTC GGC AAG GCC      51
                                                  0
Ile Ala Val Leu Thr Ser Gly Gly Asp Ala Gln G
ATC GCG GTG CTC ACC TCC GGC GGA GAT GCC CAA G                           85

INTRON 1
/gtaaggaggagagaaacaggagccatggctggtggcctggcttgctgggtcctttgcctgcaatgga    151
                                                                    Ss
gctccatctgtagcttcacactggattctcggcttcctagaaaattccttgtccgagtccccacgtt    218
tI
gtatttgttttttttttttttaatcctttttttttttttttttttaagatttatttctttatttgaa    285

agagttacacagagagagaggaaggagaggcagagagggagagaaaggtcttccatctgctggttcac    352

tcccagttggccgcagtgaccagagccatgacgatccaaagccaggagccaggggccaggaacttct    419

ctgggtctccaatgggtgcaggggcccaaggacttggcatcttctctgcttcccaggccatagtaag    486

agctgatcatagaggacagctggactcagagcgcacacaggacaagcacagcctttacgggatctta    553
                PvuII
gagtcgactacgcatgcagctgctatatgtcaatgcttcttctgtgtgaatgtacgtaatgaaacaa    620

tcgagcgagcatagttagctgtc---< 896 bp >--tagtacaaagtgtctccacgtccctggcc    1574

tgggaattttagacccaggtgtagggaagggtcctcggctctgatttccaggaaagcatgagacact    1641

cacctcaaggcaggagggtgcctcagctcaggggtgaactgaaaagaaggcagagggagggtgagca    1708

gagacaggtccctgctgtgggattgttctgccgccctgggtgttgattcacctggagacatttaagt    1775

cgtggtttccttaaatcatgacacttggacacaaacatgctgccaactgtagggtgtgtgtgcagg    1842

agagagtgtgtgtgtgtgcctcgtgtgtgctaaggtgacaggcagacctcgcaaggctagcagtgtc    1909

tgtagaagaaccctgcagccaaaatcaaggataagattgcaggcacgtccaggagacgctgcggaga    1976
          PstI
acatcgcaagcacggagtttggtttccagacattgctcagctctgctttccctaatctctctcgtg    2043

ctccttgctgtctcctcctgtaaccctgtggtccgtgttcctggccccactctcagtgacgccgtgc    2110

ctcacgcccaccccagagcacagaccaaagcctgcctctgctcgatccacccctgtcctgtgccccat    2177

cagtcacagctctccttctctcaggaactcacatcagatcacagctgcattagtgagccccttctgctc    2244
                                                  PvuII
ctgtccctagagttccagcagatctgtttgtcccaaaggccacttccagtccaggcctcctcattc    2311

cttggtctgaatgccaggagcctgcctgggcctgttccttggcgtcgtgcactgtctcccagccggct    2378

gctgatgtcttgggacacatggggctcattgtatcctgctgctcacacagttcttcagaattc----    2435
                                                                EcoRI
< 878 bp >-tatttcaatcctgagtgcctgcctccagcaaaaggctgtggggxactagtgtggac    3369

ctgggggccccctacagtgagtgcactgattgggagtgggaaagcxgtgcgaggaactccgaggacc    3436
         56
cacagccgtgactgtgtgagtcctgtgcggcttccttaggaaacctctcctcactgtccccgcax/    3502
          *
EXON 2
 28
 ly Met Asn Ala Ala Val Arg Ala Val Val Arg Val Gly Ile Phe Thr Gly
 GT ATG AAT GCT GCC GTC CGG GCT GTG GTT CGA GTT GGC ATC TTC ACT GGT      3552
                              052
Ala Arg Val Phe Phe Val His Glu
GCC CGG GTC TTC TTC GTC CAT GAG                                         3576
 SmaI

INTRON 2
/gtcagttcctcctcctcactgtgtgtccttctctgacctctctccccacccccctcctggctgcgc    3642

tgcttgctctgtctggcccccagtccccagggtcgtgtccactgtttgccagtgacgtgggaccaga    3709

ctggtttcctaccttccaaagaa-------------< 6 bp >--------------------tctxt    3743

ctgactttaaatgaagcccaacacactgacactgggggagggtgaatttctcacagggcctcttcca    3810
                                                  21
ggatcctctggtcctaaggactcagggccagtgcagcgaggctgacgggtagctatatcccacag/     3876
BamHI                                              *
EXON 3
053
Gly Tyr Gln Gly Leu Val Asp Gly Gly Asp His Ile Arg Glu Ala Thr Trp
GGC TAC CAG GGC CTG GTG GAT GGC GGG GAC CAC ATC CGG GAG GCC ACG TGG      3927
                                  078
Glu Ser Val Ser Met Met Leu Gln Leu
GAG AGC GTC TCC ATG ATG CTG CAG CTG                                     3954
                      PstIPvuII

INTRON 3
/gtaggttctggcagcaccgtcctgctctgcccgctccagtcctgccatxtggcttxagggacagga    4020

gaggagcttxtxtttgccctttatctgcgtttaaaaacagcac-----------------------    4063
                                                  96
-------< 139 bp >--------------cagctgactcagggtcagggagcttccccaggtgtc     4238
                              PvuII *
actgtgctgggagaaggggattggccatgtcagcagnggggccgccttctgcctctttcgtcaaag/    4304

EXON 4
079
Gly Gly Thr Val Ile Gly Ser Ala Arg Cys Lys Asp Phe Arg Glu Arg Glu
GGA GGC ACC GTG ATT GGA AGT GCT CGG TGC AAG GAC TTC CGG GAG CGA GAG     4355

Gly Arg Leu Arg Ala Ala His Asn Leu Val Lys Arg Gly Ile Thr Asn Leu
GGA CGT CTC CGA GCC GCC CAC AAC CTG GTG AAG CGT GGG ATC ACC AAT CTG     4406
```

```
Cys Val Ile Gly Gly Asp Gly Ser Leu Thr Gly Ala Asp Thr Phe Arg Ser
TGT GTG ATA GGC GGT GAT GGC AGC CTC ACT GGG GCT GAT ACC TTC CGT TCT   4457
                                                  1
Glu Trp Ser Asp Leu Leu Ser Asp Leu Gln Lys Ala G
GAG TGG AGC GAC CTG TTG AGT GAC CTC CAG AAA GCG G                     4494

INTRON 4
/gtgagcatggtttggacaggattatgcgtgtgcgtgtgcaaaggcaggttgtgcgcaattcgcaa   4560

tgccaatctctctgctgaatacagcacttctctctttttcatgctgga---( 225 bp )--tatt   4837

ttgaattagtaaattaaagaaagctatgacatgatacaattcaxtgtaccccggatctctaatcacct   4904
                                                                    Ps
gcaggxttcggggagcagtgtgtgccgtctcactgatcaatgctggagaaggcggggagaatgcagt   4971
tI
ccccactgattcttcccacaggcttagtcccagtctctgtccacgtgggagcaggggatcacagagg   5038

cgtgtctcccaatagctgtgtgtgtggactgggaggatctactgcccaggataaggactgcagctca   5105
                                                                  PstI
ctgcaccacgcaaaatcaggcctggctggtgctxtggcgtagtggacttcccacctgtaccgcttct   5172
                                                        PvuII
gatcagctccctgctgatggcctgggggaagcagcgxaagatggcccaggtgcttgggccctgcaccc   5239

acatgggagacctggctgaagctcctgcttcagcctggtccagccctggccattgtgaccatttggg   5306

gagtgaaccagcagatgaaagatctctctctctctctctctctctctctctctctctctctgcctct   5373

gcttttcccctctctgaaactctgactttaaaataaataaacctttaaaaaaaagaaaaaatcaggcct   5440

ttgtcccaaaaagcctggcaaatctgggtgtctggattctgtgcacagtagaagagaaacctagcag   5507

gccggcgcacagctcactaggctgaatcctccacctgtgcgccggcacactgggtttagtcccggtt   5574

gctcctcttccaggccagctctctcctgtggcccgggaxgagtggaggatggtccagctggttgggc   5641
                                  SmaI                          PvuII
cctgcacctgcatgggacaccaggxgxaagtgcctggctcctggcttcggatcggcgtgtgtgtgct   5708

ggccatatagtgcctagcggcacttgggggatacatgaaa---( 900 bp )--ctcgataccgtt   6660

gaacacgatcaaggtccgagaaaaagagacaggccccgtacccatttcttacctgaattgtgtgttct   6727

gtgttgagtcgtctcattcttctatattcatttctcaaaaaagaaatgcaggtgggcaacaaataa   6794

caaaaacatgctcggtacc----------( 405 bp )----------gagcctcac   7227
          KpnI
tgggctgataccttccgtctgagtgaagcgacctgtccagtgacctccagaaaxcggtgagacagtg   7294

gtttggaacaggattaatgcgtgtgcgtgtgcacacaggcacccttgtgccgcacattcgcacaatg   7361
                                                              72
cccacaccctcctcctgctggagatgagcaggcagcctgtgccctcctcctctgtgtccattgcc   7428
                                                              a
tgcgacgtcctggccttggctcgctccctcctagaatcggctccaggttt gcttcttcttgtcag/   7494

EXON 5
 42
 ly Lys Ile Thr Ala Glu Glu Ala Thr Arg Ser Ser Tyr Leu Asn Ile Val
 GT AAG ATC ACA GCC GAG GAG GCC ACA AGG TCC AGC TAC CTG AAC ATC GTG   7544

Gly Leu val Gly Ser Ile Asp Asn Asp Phe Cys Gly Thr Asp Met Thr Ile
GGC CTG GTC GGC TCC ATT GAC AAT GAC TTC TGC GGT ACG GAT ATG ACC ATC   7595

Gly Thr Asp Ser Ala Leu His Arg Ile Thr Glu Ile Val Asp Ala Ile Thr
GGC ACC GAC TCT GCC CTG CAC CGG ATC ACA GAG ATT GTG GAT GCC ATC ACC   7646
                19
Thr Thr Ala Gln Se
ACC ACA GCC CAG AG                                                   7660

INTRON 5
/gtacacactatgacgggcagaagagctggggcctcgtgggtgaagtgaagtctttgagctgtggtg   7726

actgatggccagtttcctggggctggtatgctatttttagctgcaagtgattcctagagcgtaggaa   7793
                                                98
tggcactggtcoctaagcactgtgagatttgccttttgccacctaccagcctcctgtgctgcatgtg   7860
                                                s
gctcctgggcccgccttgcagccatctggtgggggtggcctgtgtgcgggtgtaagtgctgtctctc   7927

tcttgcag/                                                            7935

EXON 6
  7                                                          21
  r His Gln Arg Thr Phe Val Leu Glu Val Met Gly Arg His Cys Gl
  C CAC CAG AGG ACA TTT GTG TTA GAA GTG ATG GGC CGC CAC TGT GG        7980

INTRON 6
/gtaaggtccccgtcct----------( 236 bp )----------                      8232

----------acgctatggcaagtgggtgtgagtaaggaaggtctcctcaggagccag          8280

cgatctgacctcaccctgaggtgacacacaggatctgcagaccttcccagctgtggcacccgaagag   8347
                                              26
tttgctgcgtctgggguccctgttgaggggtgcagctctgagtgggtgccctggctctctctcag/   8413
                                              s
EXON 7
  2
  y Tyr Leu Ala Leu Val Thr Ser Leu Ser Cys Gly Ala Asp Trp Val Phe
  A TAC CTG GCC CTG GTC ACC TCT CTG TCC TGC GGT GCT GAC TGG GTC TTC   8462

Ile Pro Glu Cys Pro Pro Asp Asp Asn Trp Glu Asp His Leu Cys Arg Arg
ATT CCT GAG TGT CCT CCG GAT GAC AAC TGG GAG GAT CAC CTG TGT CGC CGG   8513
       248
Leu Ser Glu
CTC AGT GAG                                                          8522

INTRON 7
/gtgctgcatgcttatccttgcaggacatgccctgtccatgtctgcaggttcatttgctcxgtctac   8588

a----------( 40 bp )----------                                       8629
            52                          23
---gagtgattaagatccaacatgctgcactggccagtcctcatgctgtttctctgcxtgaccag/   8691
           s                           s
```

```
EXON 8
249
Thr Arg Thr Arg Gly Ser Arg Leu Asn Ile Ile Ile Val Ala Glu Gly Ala
ACA CGG ACC CGT GGC TCT CGC CTC AAC ATC ATC ATT GTG GCC GAG GGT GCA   8743
                                                        280
Ile Asp Arg Asn Gly Lys Pro Ile Thr Ser Glu Gly Val Lys Asp
ATC GAC AGG AAC GGG AAA CCG ATC ACC TCA GAA GGC GTC AAG GAT           8788

INTRON 8
/gttaxtgtgcaggaagctgagggccctagcatggaggccccagaatattccctagtggttctgtg   8854

tctccccaagtccctgcctgtcattccctctgcaccacgaaccactatggacaaggggccccagcgtt   8921
                      36      27
cctggtcagccccagtcccctgactcagcagaccccctctgtgtgcttggtcccctcag/         8979
                      *       *

EXON 9
281
Leu Val Val Arg Arg Leu Gly Tyr Asp Thr Arg Val Thr Val Leu Gly His
TTG GTG GTG AGA CGT CTG GGC TAT GAC ACG CGG GTT ACC GTC CTG GGG CAT   9030
                                                    311
Val Gln Arg Gly Gly Thr Pro Ser Ala Phe Asp Arg Ile Leu
GTG CAG CGA GGT GGA ACC CCC TCA GCC TTC GAC AGG ATC CTG               9072
                                               BamHI
INTRON 9
/gtgagtcacaggctctgtggccccctggaxcatcagaattgtctgcaccttt---< 600 bp >--g  9724
                                                                    S
agctcttagcagcagtgatgttgattgagatggtgttccgtctcagccctttcccacttgcatggagg   9791
stI
ggagaggcgagtcagatggggaatttgtttctccggaaccgaggagaaccttggtttgatagggagc   9858

ggacaagtgactcagcttgaggtgtgactgtagggagggacagcaatctgctgctagccttcatgtg   9925

atacagtatgaagcagcgtgcagctctcagtggactagatgcagcactgcaatttctgctt<21bp>   10007
                          39              21
ctgcagcatctctgagtcggtttcctcatctgtaaaatgttcctgaggaggacagaggtcctgcag/   10073
PstI                      *               *                   PstI

EXON 10
312
Gly Ser Arg Met Gly Val Glu Ala Val Met Ala Leu Leu Glu Gly Thr Pro
GGC AGC AGG ATG GGC GTG GAA GCA GTG ATG GCG CTG TTG GAG GGG ACC CCG   10124

Asp Thr Pro Ala Cys Val Val Ser Leu Ser Gly Asn Gln Ala Val Arg Leu
GAC ACC CCG GCC TGT GTG GTG AGC CTG TCT GGT AAC CAG GCC GTG CGC CTG   10175
                          353
Pro Leu Met Glu Cys Val Gln Val
CCG CTC ATG GAA TGT GTC CAG GTG                                      10199

INTRON 10
/gtaagtccagcccagccctccaaaggccaaacccctctcagagcatggggagaggtxcctgcgg    10265

cxtgtcagagaagt---< 217 bp >-----ctgacacagtaattccactgttaactcatccagac   10531

agcattctgaaaatgctgatgcttgagcacgcgcctgccccagagcacagagcacagcgcagctgc   10598
                                                                PvuII

tgaaaatgccccccacccctgctgcttcctcagactcggtctctcgactgtctctatcacagaattc   10665
                                                                EcoRI
ggggaaagctgcggcagctgcttcccctgggctgtctgtggggtcttcctxattcgctccccctct   10732
              PvuII
ttattgaggccacgtgaagctgggattgggtxtggagccagggagatgcgggcacggccccctcccc   10799

gtggttctgaatagagcatcaggttgccctttctgttgctactcattactcagaaagttagggaga   10866

ataaaggtaacttctctgaaagttttggagacatcagaaaggtggagtgcacctttctcgagcaggg   10933
                                        29
acgggcaggctggtgggtggaaggcagccctgccctgagccttgccctactctgctcctcccctcag/   10999
                                        *

EXON 11
354
Thr Lys Asp Val Thr Lys Ala Met Asp Glu Lys Arg Phe Asp Glu Ala Met
ACC AAA GAT GTG ACC AAG GCC ATG GAC GAG AAG AGA TTT GAT GAA GCC ATG   11050
           37
Lys Leu Arg Gly Ar
AAG CTG AGA GGC CG                                                   11064

INTRON 11
/gtgaggaaatgaccagaccagctgacgagaaatgagatgaccaggg----------           11110
                PvuII
< 297 bp >caaagtctgcctggctccataccaaagcaggagcttctgccccctcccctttcctccc   11464

tgccacagcttctgcccagcacagatggctcagtcctctcctggccattcttctggctctaggcgag   11531
             116
gggccagctctgtcgtgatgagcaaggcccgggaggagagtgcgcagggagcatgtgaaggaggag   11598
             *          SmaI
gcccctctgcagccccgctctgcctgctgcgcaggggaatgggctgaagacacttcattttcacag/   11664
      PstI

EXON 12
  5
  g Ser Phe Met Asn Asn Trp Glu Val Tyr Lys Leu Leu Ala His Ile Arg
  G AGC TTC ATG AAC AAC TGG GAG GTG TAC AAG CTT CTG GCT CAC ATC AGA   11713
            396                     HindIII
Pro Pro Ala Pro Lys
CCC CCA GCC CCC AAG                                                  11728

INTRON 12
/gtacgacaacgggggcgcccctcactctttccctccccacccctcagtcccacctaagcctttccag  11794

cctgtgtgcaacgctgatctcagttccacccgggcggttttcttcttggccacccagatcctggccct  11861

gtgttcttagtgtgtcccggatgtctctgacccttgcccagtctgagcagcggatgcagtggacgggc  11928

aagaagtgtctgtgctttaccctctggcaaccaatcacatgcgcccgacaggagcagaggtctcgc   11995

gttggagccccgaacacacgtgaagtgatgggttggtcttgggcgagcgggctccgggggggg      12062
             83
ttgcagacagacccctcatctgcccxgtccgcagaacagggagcatgtxctgccctggtcccctgggc  12129
             *
cctcagtcctcctcccc/gactatgtagagtag/                                   12161
```

```
EXON 13
397                          ———› C-HALF
Ser Gly Ser Tyr Thr Val Ala Val Met Asn Val Gly Ala Pro Ala Ala Gly
AGT GGC TCG TAC ACC GTA GCT GTG ATG AAC GTG GGT GCC CCG GCG GCA GGC  12212

Met Asn Ala Ala Val Arg Ser Thr Val Arg Ile Gly Leu Ile Gln Gly Asn
ATG AAT GCA GCT GTT CGC TCC ACT GTG AGG ATT GGC CTG ATC CAG GGC AAC  12263
        PvuII                                                  446
Arg Val Leu Val Val His Asp Gly Phe Glu Gly Pro Ala Lys Gly Gln
CGG GTG CTG GTG GTG CAC GAT GGC TTC GAG GGG CCT GCC AAG GGC CAG      12311

INTRON 13
/gtatggaactccgagaacaaggccacatctggggccccggggctgacgtggggagactgtcctcc  12377
                                    SmaI
ctgctgatctcaggcagaagcaaatgcttgattccaggctataggaagaggcaggcgctcgggagtg  12444

aagctgttgcagatattggtcccactttgcacatggccggcaccaacattctgttttcccatgagca  12511
                                                  52
gcaaggcgcctgtgctgctctctgggccgtagctcacagggcagtcccacccaggccccctgtgctat  12578
                                                    *
gggcaccatgcagcgtgtggcctttttcctgcag/                                   12611
                          PstI

EXON 14
447
Ile Glu Glu Ala Gly Trp Ser Tyr Val Gly Gly Trp Thr Gly Gln Gly Gly
ATT GAG GAA GCC GGC TGG AGC TAC GTA GGG GGC TGG ACC GGG CAA GGC GGC  12662
                          47
Ser Lys Leu Gly Ser Lys Ar
TCC AAA CTG GGC TCA AAA AG                                          12682

INTRON 14
/gtgagcagccctgccccctcaggtgacctggtctggccccctgcagacatctgtagatggctccac  12748
                                              PstI
aggcttggcaggatcgggggtggtgtgggggagaggttgggggtgggggcaggaaagggcaccagttggt  12815
                          49
acttggaactgtagccttccccttgatttccccaaatccccatgccgctcctcctcctctccggtcc  12882
                          *
caactag/                                                              12889

EXON 15
  0
  g Thr Leu Pro Lys Lys Ser Phe Glu Gln Ile Ser Ala Asn Ile Thr Lys
  G ACT CTA CCC AAA AAG AGC TTC GAA CAG ATC AGT GCC AAC ATA ACC AAG  12938
                                                  499
Phe Asn Ile Gln Gly Leu Val Ile Ile Gly Gly Phe Glu
TTT AAC ATC CAG GGC CTG GTC ATC ATT GGG GGC TTT GAG                 12977

INTRON 15
/gtgagtgtcatcccgcccgccgctgttgggggcttgggggtgatgccccctccctgtcactgctgg  13043
          87                                  53
gggtttaggctgatgcccctccctgtcactgctgggcctxaggctgactgccccctctctgccctccc  13110
          *                                   *
cgaccccatgagtcctgctgctctcccgtcag/                                     13142

EXON 16
500
Ala Tyr Thr Gly Gly Leu Glu Leu Met Glu Gly Arg Lys Gln Phe Asp Glu
GCT TAC ACC GGG GGC CTG GAG CTG ATG GAG GGC AGG AAG CAG TTT GAC GAG  13193
                                                            Sst
Leu Cys Ile Pro Phe Val Val Ile Pro Ala Thr Val Ser Asn Asn Val Pro
CTC TGC ATC CCT TTC GTG GTC ATC CCT GCC ACA GTC TCC AAC AAC GTT CCC  13244
I                                                          550
Gly Ser Asp Phe Ser Val Gly Ala Asp Thr Ala Leu Asn Thr Ile Cys Thr
GGC TCA GAC TTC AGC GTC GGG GCG GAC ACG GCG CTC AAC ACC ATC TGC ACG  13295

INTRON 16
/gtgcagagcctgccccgccccgccccttcctgcgaggctgtgcgtgtggaccaggtgcggtcctcg  13361

ggacagctccaggcatttgagcagcgcctccctcatgagagagtgtcggagggccgcggctgttccc  13428

gccctgagcagaaaggatgggcgagttccctcccagcagagcagttcttcgtgggctgcgaacacg  13495

cctgttgtctagttgtggggaacacacggggctagcgcgctgccggccgagggagaacctgggct  13562
                          26
ccagtagctccagcagatccaaaatgactctgctttccccgtgcctgcag/                  13611
                          *                 PstI
EXON 17
551
Thr Cys Asp Arg Ile Lys Gln Ser Ala Ala Gly Thr Lys Arg Arg Val Phe
ACG TGT GAC CGT ATC AAG CAG TCC GCA GCG GGC ACC AAG CGC CGG GTG TTC  13662

Ile Ile Glu Thr Met Gly Gly Tyr Cys Gly Tyr Leu Ala Thr Met Ala Gly
ATC ATC GAG ACC ATG GGC GGC TAC TGC GGC TAC CTG GCC ACC ATG GCA GGA  13713

Leu Ala Ala Gly Ala Asp Ala Ala Tyr Ile Phe Glu Glu Pro Phe Thr Ile
CTG GCA GCC GGG GCC GAT GCT GCC TAC ATT TTT GAG GAG CCC TTC ACC ATC  13764
          605
Arg Asp Leu Gln
CGA GAC CTG CAG                                                     13776
      PstI
INTRON 17
/gtagaccagacttggccctgctacttgcctggtactgtgttgtgttgctacggtgaggtctgtctg  13842

taggctgaggagcatgtgcagcatgcagtgtctgcgtagcagattgttgtaagaaagctctcttgctg  13909

acagaacttgtcatctgatgtttagctgctgactctacgctagtctgctatattgcagctgcatgct  13976
                                                          PvuIISphI
aggcatagatagag----------< 10 bp >----------ttgtagagacagctgaga  14018
                                                          PvuII
gagctcxgtagctgcctctgcagcagctctgcacggctgccgttttgctttcaggagtgttctctc  14085
SstI          PstI
acctcctggcaggcctggaccctactttcacagagcaccacaggtccttatttttcccagatcccag  14152

tgcaaagtgtagtacaggagctgagggctctcccagtccaagagcctggccagcccttcccggagca  14219

gggggcagcatctgcacccctgcacccttgcacatccacttggcactcagaggacctcgctgcatcca  14286

ggcgcctgactgctttgcatcggcagcacagtgtgcccggatgagggaggctggaggtgggtggg  14353
                                    SmaI
```

```
cggctgaactgggggtcctgctgttgcacggagcttctggaaagggggcagagcaggccagggctgtca  14420
                                        37
gtgtgccctgtgcatcaccagctttgcccactcccagctaggcccagggttaactttggctttcag/  14486
                                        *

EXON 18
606
Ala Asn Val Glu His Leu Val Gln Lys Met Lys Thr Thr Val Lys Arg Gly
GCG AAT GTT GAA CAC CTG GTG CAA AAG ATG AAG ACG ACT GTG AAG AGA GGC  14537
          62
Leu Val Leu Ar
TTG GTG CTG AG                                                       14548

INTRON 18
/gtactgcagctgcggcttgtccgggggaagaaagggggaggcaggctttggcctacctgtggaacgg  14614
      PstIPvuII    81
gtccatggttgggggtgccctgaggaatccattgtgcatgggcgggggggggcaggcacagagcctttt  14681
                  *
ggtggcaggaagtcaggcctggccctccctggcag/                                 14716

EXON 19
  6
  g Asn Glu Lys Cys Asn Glu Asn Tyr Thr Thr Asp Phe Ile Phe Asn Leu
  G AAC GAG AAG TGC AAT GAG AAT TAC ACC ACG GAC TTC ATC TTC AAC CTG  14765

Tyr Ser Glu Glu Gly Lys Gly Ile Phe Asp Ser Arg Lys Asn Val Leu Gly
TAC TCT GAG GAG GGG AAG GGC ATC TTC GAC AGC AGG AAG AAC GTG CTT GGC  14816
          663
His Met Gln Gln
CAC ATG CAG CAG                                                      14828

INTRON 19
/gtatgcgctgcgctcgatggcagccagcagctgcaggccgtgcacatggcaggtgtcgttggtgac  14894
                                  PvuIIPstI
tgtctcgtataccagtgtcaaggaccttgcaacagaagcacctctggacatcaggaaatcgtctcag  14961

atgaggacgggggctttagagagagaaagttatttccagtgtgcaggccttctcccaccgtgctcat  15028

cggcgggagxctgctgtcttccagggagagtcccccgtccaggcaggccgtactcaxagctcccccca  15095

ggtgttctcttccaaaagtggaccctggtgtctcagctggaagggtattctgggktcaggtgcccc  15162
                                  PvuII
gacctctgagtccagcggctttcttttgatgttttacctcctcctgaggggcagcccttggagtccct  15229

tagccatccgtgctgggcagaagctgtggcagggaggaaggggagggcagaagctctgctttkggt  15296

ctgagcaggcagactttgggcttatctggtatctcatttctgtagctgtctgggatcttacatcgac  15363

tcagcatca----------( 10 bp )----------aacccactctacaag  15397

ctcttactctctctagcactgcacctxgcxagactctcttctgtxgtgctgaccaxacagcaxagcc  15464

tgcataggcagcxtggaxtgcagccccttcccttcctggttctctcctgccagggagacggccagct  15531
                                                83
ccccctggctgctgacaccatcctaggactcccttcacaggcagcatctggatttctgtgtctgccca  15598
                                                *

ctgttccctctgcttgtgtatccatgagtctctgccatccctgccacctgtgtaattcttttcag/  15664

EXON 20
664
Gly Gly Ser Pro Thr Pro Phe Asp Arg Asn Phe Ala Thr Lys Met Gly Ala
GGC GGG AGC CCC ACT CCC TTT GAC AGG AAC TTT GCT ACT AAG ATG GGA GCC  15715
                                                              6
Lys Ala Met Asn Trp Met Ala Gly Lys Ile Lys Glu Ser Tyr Arg Asn G
AAG GCC ATG AAC TGG ATG GCT GGG AAG ATC AAA GAG AGT TAC CGC AAC G    15764

INTRON 20                    42                        17
/gtaggtagggtgggagtgcgtgccctgtgcagaggctgggtccccggtgggcagctcatggtgggt  15830
                             *                         *
gattgcag/                                                            15838

EXON 21
 97
 ly Arg Ile Phe Ala Asn Thr Pro Asp Ser Gly Cys Val Leu Gly Met Arg
 GG CGG ATC TTC GCC AAC ACC CCT GAC TCC GGC TGT GTT CTG GGA ATG CGT  15888

Lys Arg Ala Leu Val Phe Gln Pro Val Thr Glu Leu Gln Asn Gln Thr Asp
AAG AGG GCT CTG GTC TTC CAA CCA GTG ACT GAG CTG CAG AAT CAG ACG GAC  15939
    73                                    PstI
Phe Gl
TTT GA                                                               15944

INTRON 21
/gtgagtgcactgctgtctccagtgtccacatcctgtcaggcacggtgctgtcctccacctggtcca  16010

ctgtcttcagtgccttttttttagaagcagtgctgtcacacttagcatttggatgtgatgccagggcc  16077
                          54                        24
ctgcagccccagcttcttcccataagctgctggtagaagctgagaaattttatcaccatctccctca  16144
PstI                      *                         *
ttcttctgtag/                                                         16155

EXON 22
  2
  u His Arg Ile Pro Lys Glu Gln Trp Trp Leu Lys Leu Arg Pro Ile Leu
  G CAC CGA ATC CCC AAG GAA CAG TGG TGG CTG AAG CTG CGG CCC ATC CTC  16204

Lys Ile Leu Ala Lys Tyr Glu Ile Asp Leu Asp Thr Ser Glu His Ala His
AAA ATC CTA GCC AAG TAC GAG ATT GAC TTG GAC ACC TCC GAG CAC GCG CAT  16255
                                                779
Leu Glu His Ile Ser Arg Lys Arg Ser Gly Glu Ala Thr Val ***
CTG GAG CAC ATC TCC CGA AAA CGG TCT GGG GAA GCC ACC GTC TAA          16300

gcctcggagcgagggacaggctgtctgctcacatcaagcctgcacgccgtgtccatgtcgtagctct  16367
                                                              ------>
ttagtcattagggtccctgtacttgtgcaccaaggccgtgaccagctctggccgggaacggagcagc  16434
        ------>                                       ------>
agtaggtggaaggctcctgtagtagactaggtcatgacccgcaccccagcctggtatgttacaaggca  16501
    Linker                                  ----------\lu consensus-
gggtcctgtagtgctctgctggacaaagtaggtccttttaaggtttatttattttgtttgtgtaagagt  16568
------>    ------>
gtcgctctcacatgcctgtcgcgaccggacgaacgagacgctcgatgtgcagagctgtgggcctcaag  16635

gtagtgaagatcgcggacgaccacaggctggctgcgtgaagtctgcagacttcacgcagccagcctg  16702
                                        PstI
tggtagtccgcgtacttcactacctctxaggcccagagctctgcacactgagccgcctcctctgtcct  16769
                                    SstI
gctcxtgcagacaaagtgggacccagtgctgtggtgcaaaggggaagctacgtgcaattgxacatt  16836
                                        ------>
tgagctgttgtcgtctcatcatcagtctgtaattgaxatagagctgca----------  16884
```

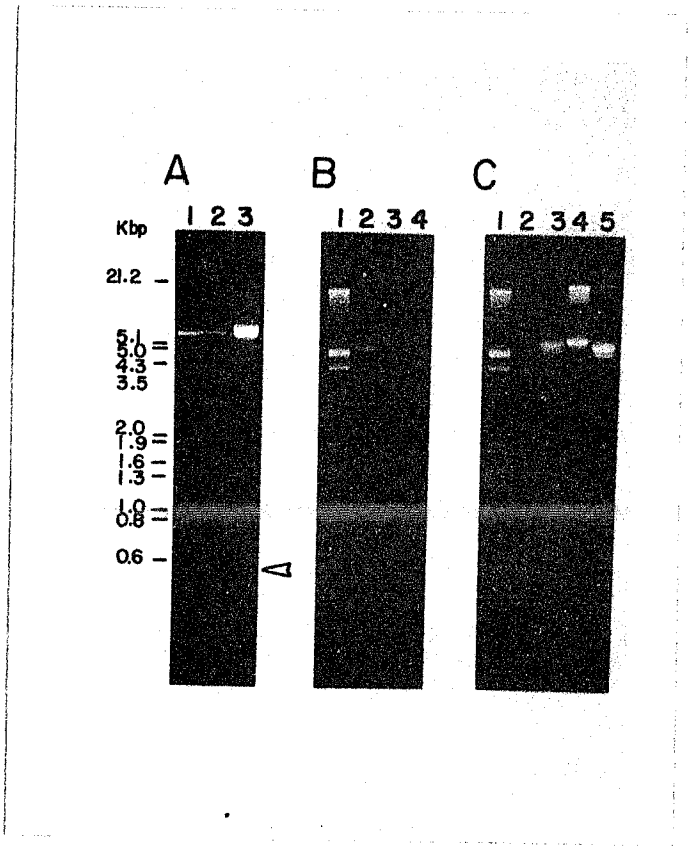

Figure 9. **Gel electrophoresis to identify M13 clones containing the N-terminal of rabbit muscle PFK gene.** A 0.5 kbp PvuII fragment selected by hybridization od lambda Charon 4A/pfk to synthetic N-probe was subcloned into the SmaI site of M13mp18, phage particles were first selected for insertion by their lac phenotypic change. Two positive clones were finally chosen and tested. (A) Mini-screen of RF DNAs. Two positive M13RF DNAs prepared as described in MATERIALS and METHODS were digested with EcoRI and HindII. The digests were subjected to electrophoresis on 0.7% agarose gel and stained by ethidium bromide. M13mp18 vector DNA was digested with the same enzymes and run on lane 3 as the control. Lane 1, clone 1; Lane 2, clone2. The open arrow indicates a 0.5 kbp fragment on lane 1 and 2 but not 3. (B) Direct gel electrophoresis. Positive M13 phages supernatant were treated as described in MATERIALS and METHODS and subjected to electrophoresis on 0.7% agarose gel and stained by ethidium bromide. HindIII and EcoRI digests of lambda DNA and M13mp18 phage supernatant treated as described in MATERIALS and METHODS were run on lane 1 and 2 as size markers. Lane 3, clone 1; lane 4, clone 2. (C) Complementary test. Positive M13 phages supernatant were treated as described in MATERIALS and METHODS. Lane 1, HindIII and EcoRI digests of lambda DNA; lane 2, M13mp18; lane 3, clone 1; lane 4, clone 1 annealed with clone 2; lane 5, clone 2.

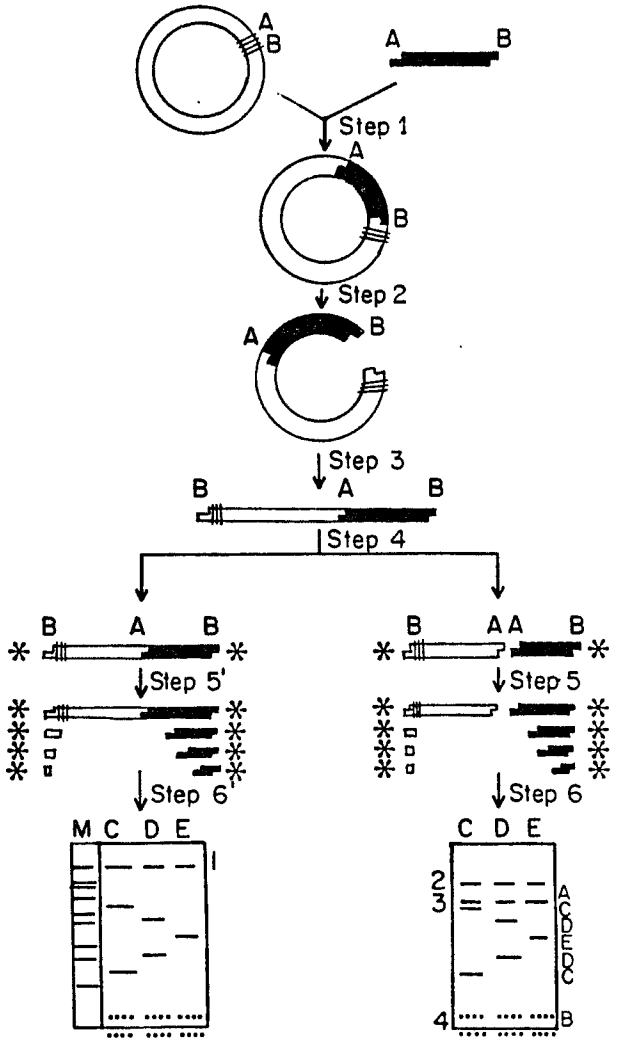

Figure 10. Schematic representation of a new strategy for rapid DNA restriction mapping.

electrophored and then gel will be dried on gel drier and exposed under an X-ray film overnight at room temperature. C, D and E represent restriction enzymes used for partial digest; M, the size markers, band 1, vector with target DNA; band 2, vector DNA only; band 3, intact target DNA; band 4 and the dotted line on the gel bottom represent the short labelled polylinker, which will run off the gel.

Figure 11 showed the gel electrophoretic patterns and the derived restriction maps of the two DNA fragments purified from RMPFK gene. The wild type lambda DNA was first used to titrate different restriction endonucleases to find the best condition for partial digestion (data not shown). The HindIII and EcoRI digests of wild type lambda DNA were labelled as the size markers at 5' end. 10^3 to 10^6 cpm was previously loaded on a 1.2% agarose gel to quantitatively check the counts needed to show clearly on autoradiogram (data not shown). The data showed that 10^4 cpm for each band was satisfactory. After exposure, the dried gel was soaked and stained by EtBr. The EtBr stained gel pattern was compared with autoradiogram to obtain a quatitative idea of how much DNA corresponded to each 10^4 cpm band. The units of restriction enzyme used for partial digestion could therefore be determined by counts of labelled DNA. For example, 35 ng corresponded to 10^4 cpm and 10^4 cpm was used for each reaction, thus 0.035 units enzyme was added to satisfy the condition of the partial digestion.On panel A, a 4.5 kbp KpnI/EcoRI

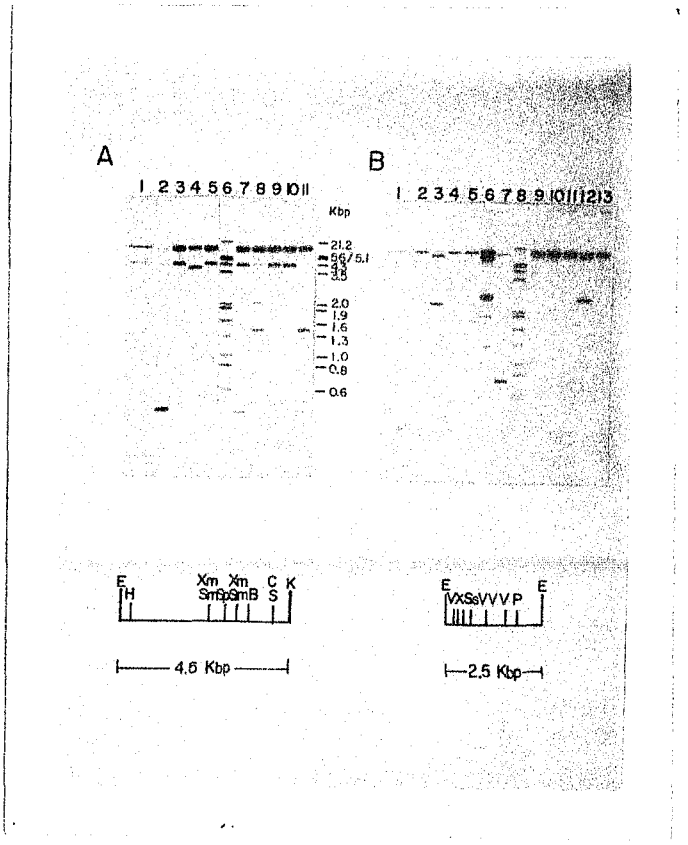

Figure 11. Rapid mapping of restriction sites in the 4.5 and 2.5 kbp DNA fragments of rabbit muscle PFK gene. All the reaction conditions were described in MATERIALS and METHODS. The principle was described in RESULTS. (A) Rapid mapping of a 4.5 kbp fragment of PFK gene. M13mp18 RF DNA containing a 4.5 kbp insert was partialy digested with restriction enzymes. The digests were subjected to electrophoresis on 1.2% agarose gel. HindIII and EcoRI digests of wild type lambda DNA was radioactively labelled at 5' end and run on lane 6 as size markers. Lane 1, intact; lane 2, SalI digest; lane 3, SstI digest; lane 4, HindIII digest; lane 5, BamHI digest; lane 7, HincII digest; lane 8, SmaI digest; lane 9, SphI digest; lane 10, XbaI digest; lane 11, XmaI digest. Deduced restriction map is given on bottom. (B) Rapid mapping of a 2.5 kbp fragment of PFK gene. M13mp19 RF DNA containing a 2.5 kbp insert was partially digested with restriction enzymes. The digests were subjected to electrophoresis on 1.2% agarose gel. Size markers was treated as described in (A) and run on lane 8. Lane 1, intact; lane 2, SalI digest; lane 3, SstI digest; lane 4, HindIII digest; lane 5, BamHI digest; lane 6, PvuII digest; lane 7, PstI digest; lane 9, HincII digest; lane 10, SmaI digest; lane 11, SphI digest; lane 12, XbaI digest; lane 13, XmaI digest. The deduced restriction maps are given on bottom.

fragment in M13mp18 was analyzed. The top band corresponding 7.2 kbp came from vector part, the second band corresponding 4.5 kbp represented the intact target DNA.By comparing the digestion pattern of XmaI (lane 11) with SmaI (lane 8) and HincII (lane 7) with SalI (lane 2), each enzyme pair showed same recognition sequence. On panel (B), a 2.5 kbp fragment in M13mp19 was analyzed. Preliminary data (not shown) revealed that target DNA did not have KpnI site. KpnI was therefore chosen to linearize the recombinant DNA. The top band represented the target DNA with vector. Its size is about 9.7 kbp. Because the complete DNA sequences and restriction map of M13mp18 and mp19 are known (Yanisch-Perron et al., 1985), some of the restriction endonucleases having restriction sites in polylinker can still be used by this strategy. For example, M13mp18 has three PvuII fragments; 6834, 322, and 93 bp. If these fragments are excluded from the autoradiogram pattern, the other fragment should come from target DNA. On panel (B) lane 6 (PvuII cut), the pattern showed the accessibility. The detailed restriction map of rabbit muscle PFK gene (Figure 12(B)) was established by using this new strategy.

3. Subcloning of PFK fragments into M13 and pBR vectors.

All the EcoRI, EcoRI/PstI, PstI, PvuII, EcoRI/PvuII fragments, were subcloned into M13 and pBR by shotgun cloning. The fragments digested by other restriction enzymes (Figure 12(B))

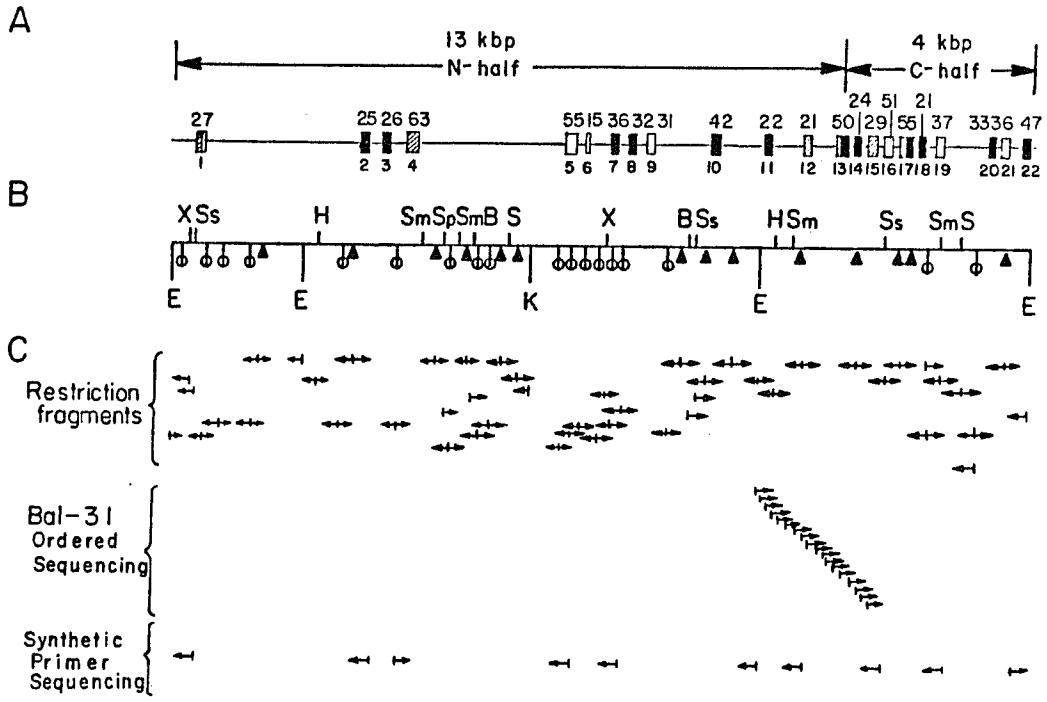

Figure 12. Genomic organization and sequencing strategy of rabbit muscle PFK gene. (A) Genomic organization. Exons are shown as boxes (see below for further definitions) drawn in approximate scale. Numbers above indicate the number of amino acid residues encoded by each exon. Exons are indicated by numbers in order from the N- to C-terminal of the encoded protein. The solid boxes indicate exons encoding nucleotide binding sites; open boxes, exons encoding sugar binding residues; crosshatched boxes, exons encoding substrate ATP binding sites; hatched boxes, exons encoding residues without assigned functions. The arrow labelled N-half spans exons encoding the amino-half of the protein while the C-half arrow spans exons encoding the carboxyl-half of the protein. (B) Restriction map Sites are: B=BamHI, E=EcoRI, H=HindIII, K=KpnI, S=SalI, Sm=SmaI, Sp=SphI, Ss=SstI, X=XbaI, solid triangles=PstI, and open circles=PvuII. (C) The sequencing strategies are grouped as restriction fragments, Bal-31 ordered sequencing, and synthetic primer sequencing.

were subcloned into M13 by forced cloning. Figure 13 shows the results of Dot hybridization to screen M13 clones generated by shotgun cloning. For example, 8.9 kbp EcoRI fragment digested with PstI were subcloned into M13mp18. One hundred and eighty seven clones were chosen and spot on nitrocellulose paper. The hybridization results were shown in Figure 13. The strength of signal roughly reflected the size of insert. Figure 14 shows the hybridization results of pBR plasmid clones. As shown in panel (A) and (B), 3 fragments: 2.5 kbp EcoRI, 1.8 kbp and 0.7 kbp EcoRI/PstI fragments were successfully subcloned into pBR3-25. This result also showed that 2.5 kbp had only one PstI site. Panel (C) and (D) showed that the fragments of 6.0 kbp EcoRI, 0.5 kbp EcoRI/PstI, 1.8 kbp PstI were subcloned into pBR325. This result also showed that the 6.0 kbp EcoRI fragment had at least 5 PstI sites. Panel (E) and (F) showed that the fragments of 8.9 kbp EcoRI, 5.0 kbp and 4.0 kbp EcoRI/SalI, 1.8, 0.5, 1.1, 0.9, 1.3, 0.3 kbp PstI fragments were subcloned into pBR325. These results also showed that 8.9 kbp fragment had at least 8 PstI sites.

4. Probing of exon number and exon positions of rabbit muscle PFK gene.

In order to probe the exon number and locate their positions, synthetic oligonucleotides were made as described (Matteucci & Caruthers, 1981). The DNA sequence and correspondant PFK posi-

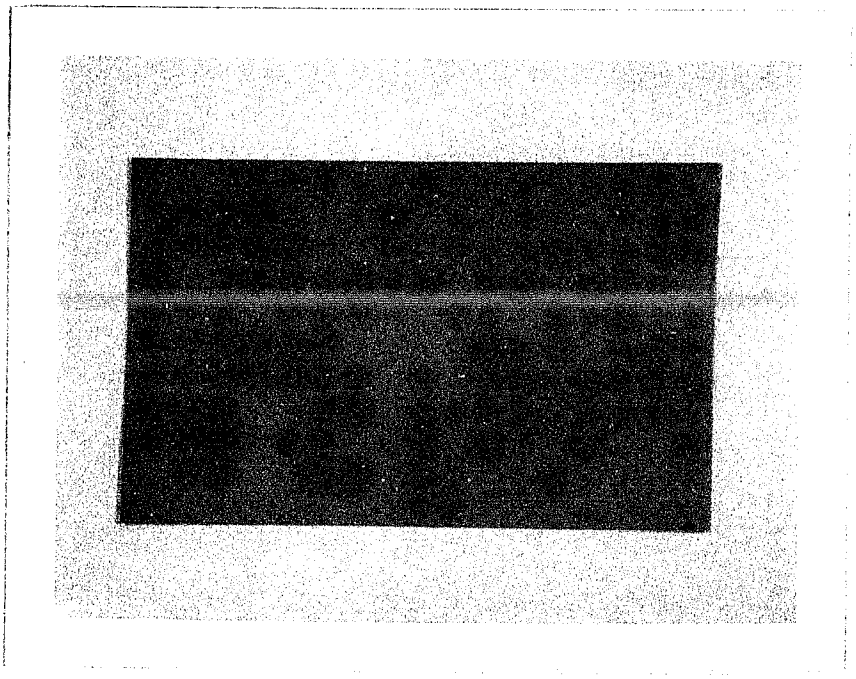

Figure 13. **Dot hybridization to screen M13 clones generated by shotgun cloning of rabbit muscle PFK gene fragments.** A 8.9 kbp DNA was purified from pBR325/8.9 kbp clone and digested with PvuII. The digest mixture was subcloned into the SmaI site fo M13mp18. Phage particles were selected for insertion by their phenotypic change. One hundred and eighty seven putatively positive clones were chosed, spot on nitrocelulose paper, and hybridized with the nick-translated 8.9 kbp probe as described in MATERIALS and METHODS.

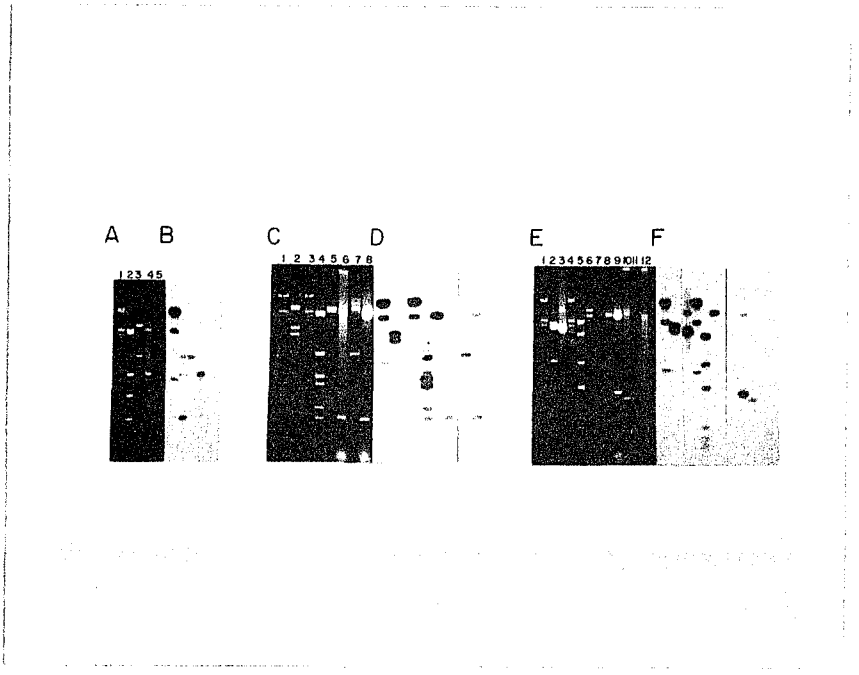

Figure 14. **Southern transfer and hybridization to identify pBR clones containing fragments of rabbit muscle PFK gene.** DNA fragments generated by EcoRI, EcoRI and PstI, and PstI digests were subcloned into plasmid 325. Plasmid DNAs with inserts selected by their antibiotic phenotypic change and toothpicker method (Barnes, 1977) were purified, digested with restriction enzymes as described in MATERIALS and METHODS. A, C, and E were the digestion pattern stained by ethidium bromide. Plasmid DNAs containing inserts from 2.5 kbp EcoRI fragment were subjected to A, from 8.9 kbp EcoRI fragment were subjected to C, and from 6.0 kbp EcoRI fragment were subjected to E. Size markers were run on lane 1 on A, lane 1 and 3 on C, and lane 1 and lane 4 on E. (A) lane 1, size markers; 2 and 3, pBR325/2.5 kbp, 2 was digested with EcoRI and PstI, 3 was digested with EcoRI; 4, pBR325/1.8 kbp, digested with EcoRI and PstI; 5, pBR325/0.7 kbp, digested with EcoRI and PstI (C) lane 1 and 3, size markers; 2, 4, and 5, pBR325/6.0 kbp, 2 was digested with EcoRI and SstI, 4 was digested with EcoRI and PstI, 5 was digested with EcoRI. 6 and 8, pBR325/0.5 kbp digested with PstI and EcoRI; 7, pBR325/1.8 kbp, digested with PstI. (E) Lane 1 and 4, size markers; 2, pBR325/5 kbp, digested with EcoRI and SalI; 3, pBR325/4 kbp, digested with EcoRI and SalI; 5 and 6, pBR325/8.9 kbp, 5 was digested with EcoRI and PstI, 6 was digested with EcoRI; 7, pBR325/1.8 kbp, digested with PstI; 8, pBR325/0.4 kbp, digested with PstI; 9, pBR325/1.1 kbp, digested with PstI; 10, pBR325/0.9 kbp, digested with PstI; 11, pBR325/1.2 kbp, digested with PstI; 12, pBR325/0.3 kbp, digested with PstI. Gels were blotted, hybridized with nick-translated lambda Charon 4A/pfk probe as described in MATERIALS and METHODS. B, D, and F indicating the hybridization results were arranged site by site for each gel.

tion were shown in Figure 15. To clarify the digestion pattern, the gel was blotted and hybridized with vectors. In Figure 16(A), the blot was hybridized with the nick-translated wild type lambda DNA; in (C) the blot was hybridized with the nick-translated pBR325 DNA. From the results of panel (A), we excluded bands from lambda 4A. From the results of panel (B), we excluded bands from pBR325. It is interesting to note that among these DNAs, there is no significant homology. Therefore, it is convenient to use the nick-translated lambda 4A/pfk DNA as probe to screen any library containing the PFK gene fragment, and to avoid the tedious steps to isolate pure fragment from gel. In Figure 17, multiple blots were made and hybridized with oligonucleotide probes which corresponded to different regions of RMPFK. These oligonucleotides also represent different exons. EcoRI/PstI or EcoRI/PvuII digested lambda 4A/pfk DNA were loaded on lanes 1 and 7 as the positive control. The results of PFK1 probe on panel (B) showed a 1.7 kbp PstI and a 0.5 kbp PvuII bands. This result reveals the location of the first exon. The results of PFK1.5 probe on panel (C) showed a 1.8 kbp PstI band. In the same way, a 3.0 kbp PstI band hybridized with PFK 2 probe; a 1.2 kbp PstI band hybridized with PFK 4 probe; a 0.5 kbp PstI/EcoRI band hybridized with PFK 7 probe. It was demonstrated again that the 0.5 kbp PvuII fragment contained the N-terminal of the gene; while the 0.5 kbp Pst/EcoRI fragment contained the C-terminal of the gene. In this way, 10 exons

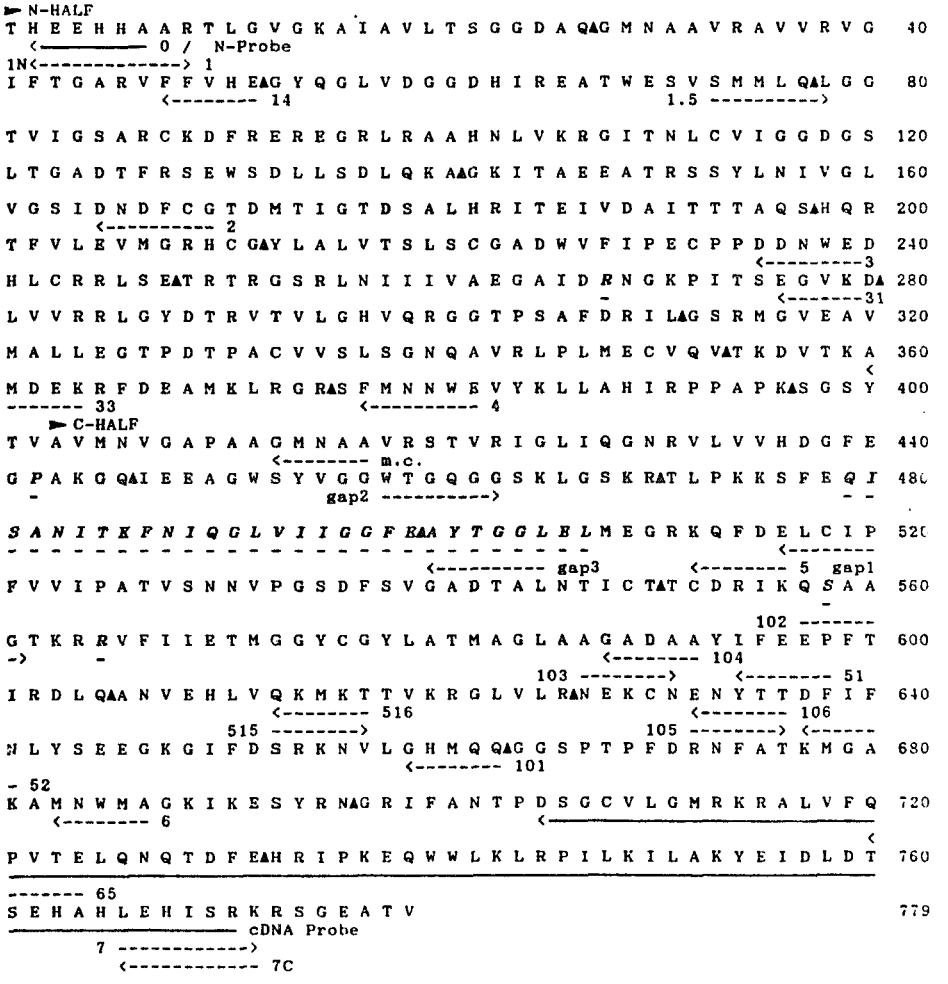

Figure 15. Corresponding positions of synthetic oligonucleotides and cDNA probe in RMPFK. The sequence for each synthetic oligonucleotide was described in MATERIALS and METHODS. The number in the right margin correspond to the sequence from the first amino acid residue (number with 1) to the last amino acid residue (number with 779) of rabbit muscle PFK. The arrow indicates the 3' end of each DNA sequence. The solid line under the sequence indicates the corresponding position of cDNA probe and N-probe used to determine C- and N-terminal of PFK gene. The solid arrow head indicates the first amino acid of N- and C-halves of the protein. The solid triangles represent the introns. The italic letters represent corrected amino acids or newly identified amino acids.

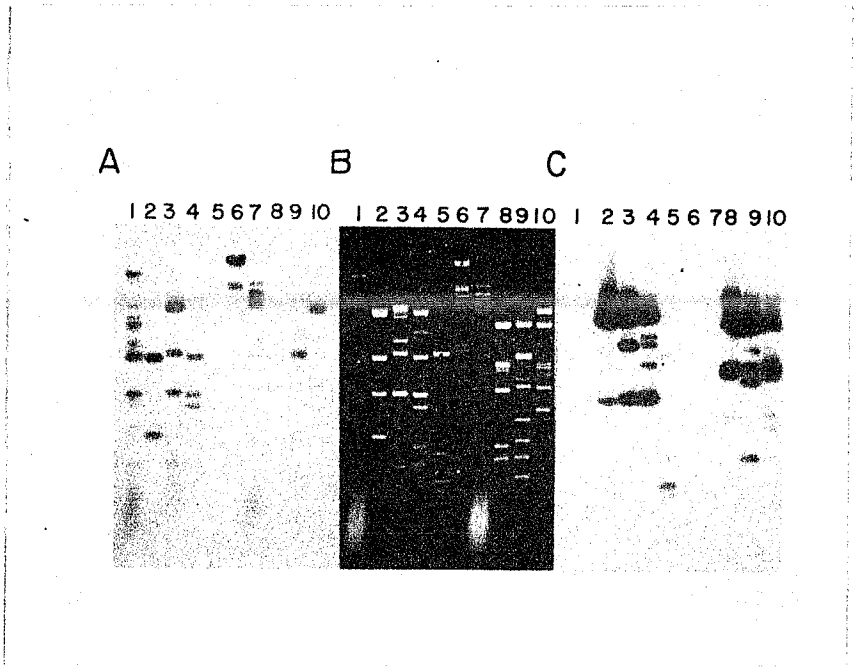

Figure 16. Southern transfer and hybridization using nick-translated vector DNAs. 3 EcoRI fragments of rabbit muscle PFK gene; 2.5 kbp, 8.9 kbp, and 6.0 kbp, were subcloned into the EcoRI site of pBR325 respectively. Plasmid pBR325 DNAs with different insert were purified as described in MATERIALS and METHODS and digested with EcoRI and PstI or EcoRI and PvuII. The digests were electrophoresed on a 1.2% agarose gel, blotted and hybridized with nick-translated pBR325 DNA oe wild type lambda phage DNA. (B) Gel electrophoresis of digests, lanes 1 to 4 represent the PstI digest, whereas lanes 7 to 10 are PvuII digest. HaeII digest of pBR322 DNA and HindIII and EcoRI digests of lambda DNA were run on lanes 5 and 6 as size markers. Lanes 1 and 7, lambda Charon 4A/pfk; lanes 2 and 8, pBR325/2.5 kbp; lanes 3 and 9, pBR325/8.9 kbp; lanes 4 and 10, pBR325/6.0 kbp. (A) Blot was hybridized with nick-translated wild type lambda phage DNA, (C) Blot was hybridized with nick-translated pBR325 DNA.

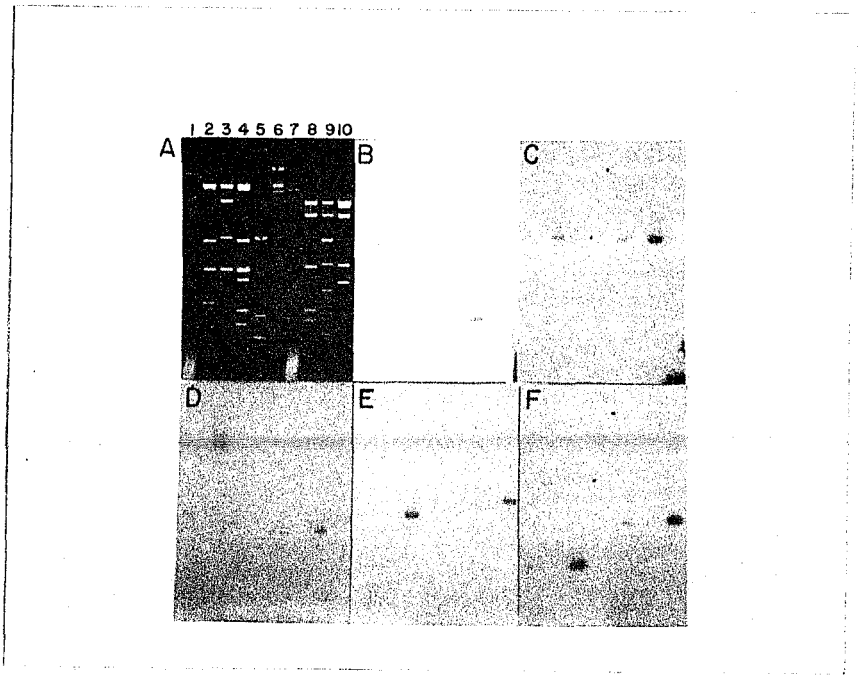

Figure 17. Southern transfer and hybridization to probe the exons numbers and positions using synthetic oligonucleotides. Oligonucleotides were synthesized and purified as described in MATERIALS and METHODS. DNA sequence and corresponding protein position were also as described in APPENDIX 4. Oligonucleotides were radioactively labelled at 5' end with T_4 DNA Kinase. (A) Gel electrophoresis pattern as described in Figure 16(B). (B) Blot was hybridized with PFK 1. (C) Blot was hybridized with PFK 1.5. (D) Blot was hybridized with PFK 2. (E) Blot was hybridized with PFK 4. (F) Blot was hybridized with PFK 7.

were found and localized unambiguously along the restriction map of PFK gene.

5. Sequence analysis of rabbit muscle PFK gene.

As shown in Figure 8 and Table 2, 71% of the DNA sequence of rabbit muscle PFK gene has been determined. These data include the sequence of all exons, 67% of the sequence of introns, 61 bp in the 5' flanking region, and 584 bp in the 3' flanking region. The promotor, cap site and polyadenylation signal of this gene are not yet identified. The DNA sequence shows that 7 bPs upstream of the translation initiation codon ATG is an C/T rich region followed by a sequence of TAG, a typical 3' splicing junction sequence. This result indicates there is an additional intron in the 5' untranslated region of the PFK gene (Chung et al., 1985). An Alu consensus sequence (CCAGCCTGG) was found at the 3' flanking region (nucleotides 16479-16487), while 3 pairs of direct repeats with sizes 155, 394, and 130 bp respectively were found (Figure 8). These data indicate a Alu-like sequence at the 3' flanking region (Schmid & Jelinek, 1982). No tRNA gene or KpnI short repeats were found in both coding and noncoding region (data not shown), however, a variety of short repeating sequences were noted. These include $(TC)_{19}$ at nucleotides 5329-5367 and a nearly continuous T track at nucleotides 222-282 (Figure 8).

The consensus of splicing junction (Breathnach & Chambon,

Table 2. Exons and introns of rabbit muscle PFK gene.

Exon	Length b.p.	Residue number encoded	Intron	Length b.p.
E 1	82	27	I 0	88
E 2	74	25	I 1	3500
E 3	78	26	I 2	300
E 4	190	63	I 3	350
E 5	166	55	I 4	3000
E 6	45	15	I 5	275
E 7	109	36	I 6	500
E 8	96	32	I 7	170
E 9	93	31	I 8	191
E10	126	42	I 9	1000
E11	65	22	I10	800
E12	64	21	I11	600
E13	150	50	I12	433
E14	71	24	I13	300
E15	88	29	I14	207
E16	153	51	I15	165
E17	165	55	I16	316
E18	62	21	I17	740
E19	112	37	I18	168
E20	100	33	I19	836
E21	106	36	I20	73
E22	142	47	I21	210
Total 22 exons	2337 b.p.	779 A.A.s	22 introns	14782 b.p.

1981) are obeyed at the 5' and 3' termini of each intron (Table 3). A second consensus sequence near the 3' end of each intron (Keller & Noon, 1984; Ruskin et al., 1984) complementary to its own 5' end and considered to be the sequence of branch site was also present (Table 4). These sequences are the ones that best satisfy the criteria of complementarity and location (within 70 bases of the 3' end), with the exceptions of intron 3, 5, and 11. The possible Lariat branch site in each intron is not unique; for example, two possible branch sites were found in introns 7, 8, 9, 15, 20, and 21. These branch sites were also indicated in Figure 8 by asterisks.

The four bases are not in equal molar amounts for both overall and coding sequences as shown in Table 5. The coding sequence has an even higher C+G content (59%), while the overall sequence has 57% in G+C content. The result of high G+C content matches the results of the high PstI and PvuII sites observed in Figure 12(B).

The results of dinucleotide percentages in the gene are shown in Table 6. Considering the high G+C content, we should not be surprised to see the low percentages of AT, AA, TA, and TT. However, the deficit of dinucleotide CG, a well established characteristic of eukaryotic genes (Nussinov, 1981), is evident in this gene. The ACC triplet coding for the first residue; theronine, immediately follows the ATG codon. The termination codon TAA immediately follows the last valine codon GTC. A

Table 3. Nucleotide sequences present at the splice junctions of the rabbit muscle PFK gene. The 5' and 3' terminal sequences of each intron are shown between the two arrows that define the splicing sites; they are flanked by the terminal sequence of the respective exons. The underlined sequences agree with the consensus (Breathnach & Chambon, 1981).

5'	GAG GTGGCT	----------	I 0	----------	TCTTGCCTTCTAGAG T 3'
	AAG GTAAGG	----------	I 1	----------	TCACTGTCCCCGCAG G
	GAG GTCAGT	----------	I 2	----------	AGCTATATCCCACAG G
	CTG GTAGGT	----------	I 3	----------	CCTCTTTCGTCAAAG G
	CGG GTGAGA	----------	I 4	----------	GCTTCTTCTTGTCAG G
	GAG GTACAC	----------	I 5	----------	GTCTCTCTCTTGCAG C
	TGG GTAAGG	----------	I 6	----------	CTGGCTCTCTCTCAG A
	GAG GTGCTG	----------	I 7	----------	TCTCTGCXTGACCAG A
	GAT GTTAXT	----------	I 8	----------	GCTTGGTCCCCTCAG T
	CTG GTGAGT	----------	I 9	----------	ACAGAGGTCCTGCAG G
	GTG GTAAGT	----------	I10	----------	CTGCTCCTCCCTCAG A
	CCG GTGAGG	----------	I11	----------	ACTTCATTTTCACAG G
	AAG GTACGA	----------	I12	----------	GACTATGTAGAGTAG A
	CAG GTATGG	----------	I13	----------	GGCCTTTTCCTGCAG A
	AAG GTGAGC	----------	I14	----------	CTCCGGTCCAACTAG G
	GAG GTGAGT	----------	I15	----------	CTGCTCTCCCGTCAG G
	ACG GTGCAG	----------	I16	----------	TCCCCGTGCCTGCAG A
	CAG GTAGAC	----------	I17	----------	AACTTTGGCTTTCAG G
	GAG GTACTG	----------	I18	----------	GGCCCTCCCTGGCAG G
	CAG GTATGC	----------	I19	----------	GTAATTCTTTTTCAG G
	ACG GTAGGT	----------	I20	----------	GGTGGGTGATTGCAG G
	TGA GTGAGT	----------	I21	----------	CTCATTCTTCTGTAG G
	Consensus Sequence				
	C A				TTTTTTTTTTT T
5'	AAG GTGAGT	----------	IVS	----------	CCCCCCCCCCCNCAG G 3'

Table 4. Possible Lariat branch sites in the rabbit muscle PFK gene. The top row of each set of paired sequences represent the 5' splice junction. The arrow indicates the splicing junction. The complementary sequence (shown on the second or third row) occurs near the 3' end of the same intron; the exact distance from the 3' splice site is indicated by the number below the base A which is the potential Lariat site. Underlined sequences indicate complementarity (column 2) or agree with the consensus sequences (column 4 and 5) including G-T wobble pairs.

Intron	Sequence at Branch Site	Base Pairing	Keller & Noon Consensus CTGAC	Ruskin et al. Consensus T T G T CNCTAAC
I 0	3' GATCTCCGTG GAG 5' 5' CTGTCCTAAC TGA 3' 26	7/13	CTAAC	TCCTAAC
I 1	3' ACGAGGAATG GAA 5' 5' CAGCCGTGAC TGT 3' 56	6/13	GTGAC	CCGTGAC
I 2	3' CCCTTGACTG GAG 5' 5' CGAGGCTGAC GGG 3' 21	8/13	CTGAC	GGCTGAC
I 3	3' GTCTTGGATG GTC 5' 5' XXCAGCTGAC TCA 3' 96	7/13	CTGAC	AGCTGAC
I 4	3' GTACAGAGTG GGC 5' 5' TCTGTGTCAT TGC 3' 72	9/13	GTCAT	GTGTCAT
I 5	3' ATCACAGATG GAG 5' 5' TGCCACCTAC CAG 3' 98	6/13	GCTAC	CACCTAC
I 6	3' CCCTGGAATG GGT 5' 5' CAGCTCTGAG TGG 3' 26	6/13	CTGAG	CTCTGAG
I 7	3' GTACGTGCTG GAG 5' 5' TGATTAAGAT GCA 3' 52	7/13	AAGAT	TTAAGAT
	5' CAGTGCTCAT GCT 3' 23	7/13	CTCAT	TCCTCAT
I 8	3' CGTGTTAATTG TAG 5' 5' TGCGCTTGAC TCA 3' 36	6/13	TTGAC	CCTTGAC
	5' CTCAGCAGAC CCC 3' 27	6/13	CAGAC	AGCAGAC
I 9	3' ACACTGAGTG GTC 5' 5' GTTTCCTCAT CTG 3' 39	8/13	CTCAT	TCCTCAT
	5' TGTTCCTGAG GAG 3' 21	8/13	CTGAG	TCCTGAG
I10	3' GACCTGAATG GTG 5' 5' CTGCCCTGAG GCT 3' 29	8/13	CTGAG	CCCTGAG
I11	3' TAAAGGAGTG GCC 5' 5' CTGTCGTGAT GAG 3' 116	7/13	GTGAT	TCGTGAT
I12	3' AACAAGCATG GAA 5' 5' AGACCCTCAT CTG 3' 83	5/13	CTCAT	CCCTCAT
I13	3' TCAAGGTATG GAC 5' 5' GCAGTCCCAC CCA 3' 52	6/13	CCCAC	GTCCCAC
I14	3' CCGACGAGTG GAA 5' 5' TTCCCTTGAT TTC 3' 49	7/13	TTGAT	CCTTGAT
I15	3' ACTGTGAGTG GAG 5' 5' TTAGGCTGAT GCC 3' 87	8/13	CTGAT	GGCTGAT
	5' TAAGCCTGAC TCC 3' 53	9/13	CTGAC	GGCTGAT
I16	3' CCGAGACGTG GCA 5' 5' GATCCAAAAT GAC 3' 26	5/13	AAAAT	CCAAAAT
I17	3' AGAACAGATG GAC 5' 5' CTTTGCCCAC TCC 3' 37	8/13	CCCAC	TCCCCAC
I18	3' CGACGTCATG GAG 5' 5' GTGCCCTGAG GAA 3' 81	4/13	CTGAG	CCCTGAG
I19	3' GTCCCGTATG GAC 5' 5' GCATCTGGAT TTC 3' 83	7/13	TGGAT	TCTGGAT
I20	3' GGGATGGATG GCA 5' 5' CTATGCAGAG GCT 3' 42	7/13	CAGAG	TGCAGAG
	5' GGCAGCTCAT GGT 3' 17	8/13	CTCAT	AGCTCAT
I21	3' CACGTGAGTG AGT 5' 5' CTTGCCATAA GCT 3' 64	6/13	CATAA	CCCATAA
	5' ATTTTATCAC CAT 3' 24	6/13	ATCAC	TTATCAC

Table 5. Nucleotide percentages in rabbit muscle PFK gene

	A	C	G	T	C + G	A + T
overall	*20*	*28*	*29*	*23*	*57*	*43*
coding	22	27	32	19	59	41

Table 6. Dinucleotide percentages in rabbit muscle PFK gene.

Italic numbers represent the percentage for the overall PFK gene, regular numbers represent the percentage for the coding sequence of PFK gene.

5' \ 3'	A	C	G	T
A	*4.6* --- 3.8	*6.2* --- 4.6	*7.0* --- 7.8	*4.3* --- 3.6
C	*7.8* --- 7.5	*8.0* --- 8.6	*5.2* --- 3.3	*6.5* --- 8.9
G	*8.2* --- 6.3	*7.9* --- 8.1	*10.6* --- 8.7	*5.0* --- 5.8
T	*1.5* --- 2.3	*5.4* --- 6.9	*8.8* --- 9.1	*3.0* --- 4.7

similar arrangement was reported for the cDNA of rabbit muscle creatine phosphokinase (Putney et al., 1984).

The amino acid sequence of the rabbit muscle PFK was revised and completed from the coding sequence as shown in Figure 15. The first ATG codon triplet codes for the initiation methionine residue which is post-translationally removed. The revised PFK has a calculated molecular mass of 84975 daltons and the composition of 779 amino acids. Our amino acid sequence agrees with that reported by Poorman (Poorman et al., 1984) except for four positions of discrepancy (Figure 15): an arginine instead of a serine at amino acid 268 (number the first amino acid Thr with 1 and the last amino acid Val with 779) a proline instead of a leucine at amino acid 442, a serine instead of an isoleucine at amino acid 558 and an additional arginine at amino acid 565. The amino acid composition is shown on Table 7. The high C + G content (~60%) of coding sequence reflects that the predominant codons end in C or G for amino acids Ala, Arg, Asn, Asp, Pro, Ser and Thr (Table 8). This result is consistent with the characteristic prevalence of G+C in many mammalian genes (Nussinov, 1981).

The 3' portion (nucleotides 12915-12977) of exon 15 and the 5' region (nucleotides 13143-13166) of exon 16 together code for an amino acid sequence of 29 residues (Figure 15, amino acids 479-507). The helix-probability and hydrophobicity profiles of these 29 amino acid residues are shown in the hatched areas of

Table 7. Amino acid compositions and percentages of rabbit muscle PFK

AA	Number	Num%	Wt%
Ala	65	8.3	5.4
Arg	54	6.9	9.9
Asn	30	3.9	4.0
Asp	39	5.0	5.3
Cys	15	1.9	1.8
Gln	24	3.1	3.6
Glu	47	6.0	7.1
Gly	82	10.5	5.5
His	19	2.4	3.1
Ile	46	5.9	6.1
Leu	60	7.7	8.0
Lys	38	4.9	5.7
Met	22	2.8	3.4
Phe	29	3.7	5.0
Pro	24	3.1	2.7
Ser	43	5.5	4.4
Thr	56	7.2	6.7
Trp	10	1.3	2.2
Tyr	15	1.9	2.9
Val	61	7.8	7.1

Length of amino acid sequence is 779.

Molecular weight is 84975.

Table 8. Codon usage in the coding sequence of rabbit muscle PFK gene

	N	C	Total		N	C	Total		N	C	Total		N	C	Total
Ala GCA	3	5	8	Arg CGA	4	3	7	Asn AAT	3	5	8	Asp GAT	11	2	13
T	6	7	13	T	5	2	7	C	8	14	22	C	14	12	26
G	3	6	9	G	10	5	15								
C	21	14	35	C	4	3	7								
				AGA	5	1	6								
				G	5	7	12								
Cys TGT	6	2	8	Gln CAA	1	3	4	Glu GAA	6	5	11	Gly GGA	6	3	9
C	3	4	7	G	8	12	20	G	17	19	36	T	9	1	10
												G	7	12	19
												C	20	24	44
His CAT	4	1	5	Ile ATA	1	1	2	Leu TTA	1	0	1	Lys AAA	3	6	9
C	8	6	14	T	5	5	10	G	3	2	5	G	10	19	29
				C	15	19	34	CTA	0	2	2				
								T	1	1	2				
								G	22	18	40				
								C	7	3	10				
Met ATG	11	11	22	Phe TTT	2	7	9	Pro CCA	1	1	2	Ser TCA	2	2	4
				C	9	11	20	T	2	4	6	T	5	2	7
								G	5	1	6	G	1	0	1
								C	3	7	10	C	5	7	12
												AGT	4	2	6
												C	8	5	13
Thr ACA	6	1	7	Trp TGG	5	5	10	Tyr TAT	1	0	1	Val GTA	1	1	2
T	2	6	8					C	5	9	14	T	3	4	7
G	3	6	9									G	22	13	35
C	20	12	32									C	11	6	17

Figure 18. A similar pattern of one alpha-helix followed by a ß-sheet observed in BsPFK was demonstrated. The high hydrophobic property of these 29 residues (as shown in Figure 18(B)) agrees with Kemp's observation (Kemp, personal communication). Furthermore, an oligonucleotide probe; gap 3 (Figure 15), was shown to hybridize to the northern blot of rabbit muscle mRNAs (Valdez, personal communication). These data show that these 29 amino acids residues fit the undetermined gap (Poorman et al., 1984). Thus, our nucleotide sequence for this gene has completed the entire amino acid sequence of rabbit muscle PFK.

More importantly, the identification of a phenylalanine at amino acid 498 and a glutamic acid at amino acid 499 excludes the C-half of this enzyme from a role in binding of substrate ATP (Poorman et al., 1984; Hellinga & Evan, 1985).

6. Genomic organization of rabbit muscle PFK gene.

The rabbit muscle PFK gene in lambda phage clone is 17 kilobase pairs in length, within the two flanking EcoRI cloning linkers (Figure 12(C)). There are 18 PvuII sites and 14 PstI sites but only two EcoRI sites and two HindIII sites (Figure 12(C)). The exon number and exon positions of this gene as shown in Figure 12(B) were established by comparing its DNA sequence with the amino acid sequence (Poorman et al., 1984) and localizing the fragments along the restriction map by hybridization. The size of the exons ranges from 15 codons (exon 6) to

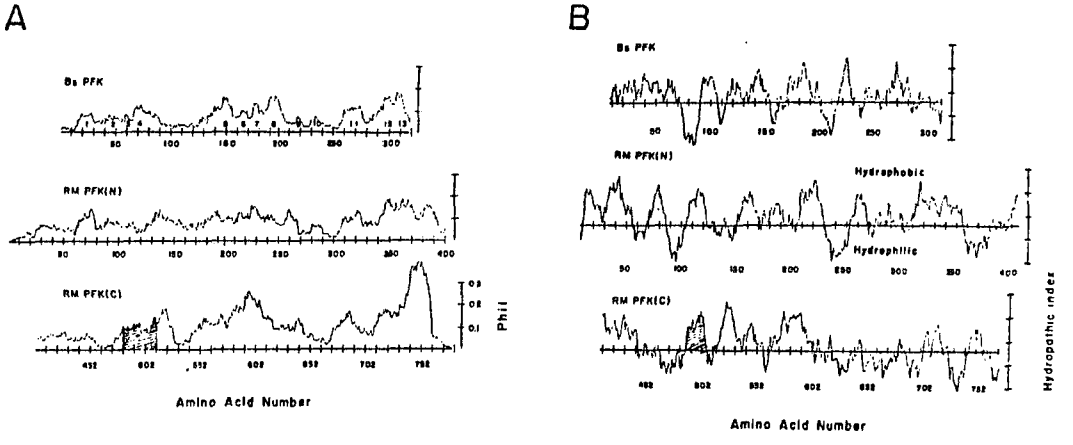

Figure 18. Comparison of the helix-probability and the hydrophobicity patterns in PFKs. (A) Comparison of the helix-probability pattern. The plots are based on the sequence presented on Figure 3. Helix-probability (Phil) are taken from Mattice et al. (Mattice et al., 1980) and averaged over a length of 10 residues. The representation of BsPFK, RMPFK (N), and RMPFK (C) are as same as in Figure 3. The positions of 13 alpha-helix structural elements of BsPFK are also indicated by numbers. (B) Comparison of the hydrophobicity pattern. The plots are based on the sequence presented on Figure 3. Hydropathic index are taken from Kyte and Doolittle (Kyte & Doolittle, 1982) and averaged over a length of 10 residues.

63 codons (exon 4). These exons can be grouped into three classes by size: 10 "50" bp; 11 "140" bp and 1 "200" bp. The abundant classes are 50 bP and 140 bp sequences. These results agree with the statistics of higher eukaryotic genes (Naora & Deacon, 1982). Approximately, 86% of the cloned gene is occupied by introns. Individual introns vary considerably in size from 73 to 3500 bp, although short introns ranging from 200 to 400 bp in length are more abundant (Table 2). The map in Figure 12(A) clearly shows that introns in the N-half region are much longer (up to 3500 bp) than those in the C-half region (73 to 836 bp). Thus, the 12 exons encoding the N-half of the protein are scattered over 13 kbp while the ten exons encoding the C-half of the protein are clustered in a length of 4 kbp.

The majority of the introns are positioned between or near the ends of the secondary structural elements as shown in Figure 19(A). It is also noted that the coding sequences for alpha-helices 6, 8, 9, and 11 are interrupted by introns. Based on Poorman's model (Poorman et al., 1984) that the RMPFK subunit resembles a head-to-tail linked dimer of BsPFK, a hypothetical tertiary structure for RMPFK is shown in Figure 19(B). When the sites of introns shown in Figure 19(A) are overlaid on this model, the image shows that the positions of these introns are on the surface of the N- and C-domains of the protein. Figure 3 shows four regions in the rabbit muscle PFK sequence standing out as "extra" residues (italicized) when the bacterial PFKs and

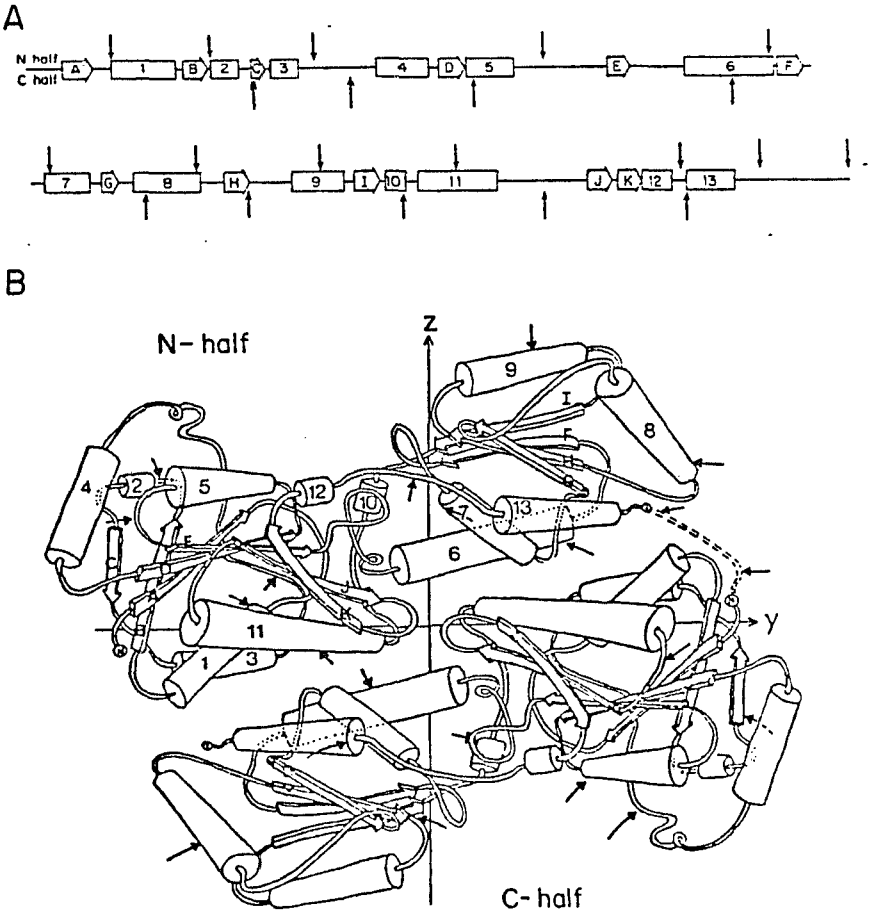

Figure 19. **The positions of introns in the structure of rabbit muscle PFK.** A. Position of introns along the secondary structural elements based on the structure of B. stearothermophilus PFK. Arrows located above indicate position of introns in the N-half of rabbit muscle PFK gene, Arrows pointing from the bottom are the C-half introns. B. Projection of intron positions on the three-dimensional structure of rabbit muscle PFK)based on the crystallographic structure of B. stearothermophilus PFK (Evans and Hudson, 1979) and amino acid sequence homology between these two proteins). The schematic drawing is viewed along the x-axis. The large open arrows represent ß-sheet strands (A-K) and cylinders represent alpha-helices (1-13). The N- and C-halves are joined by a connecting peptide (dashed lines). The small arrows point to the positions of introns in the gene sequence (modified from Evans and Hudson, 1979 with permission)

the two halves of rabbit muscle PFK are aligned for sequence homologies. Introns interrupt the coding sequences in each of these four regions. It seems likely that these "extra" amino acid residues were inserted by intron slide shuffling as proposed by Gilbert (Gilbert et al., 1986).

II. Techniques developed in this work

1. Making of strand and length specific probes.

A new technique was developed to make the strand and length specific probes. Figure 20(A) shows the restriction map of 6.0 kbp EcoRI fragment of RMPFK gene. It also indicates the position of 2.5 kbp fragment. On panel (B), lane (2) and (3) showed that circular vector was linearized but the reaction was not completed, lane (5) showed that a band migrated to the position between 1.6 kbp and 2.0 kbp. Although the digested lambda DNA was run on the same gel as markers, the corresponding size might not reflect the real size of single stranded DNA fragment. When the gel was blotted and hybridized with a nick-translated 2.5 kbp probe, it showed that the band came from a 2.5 kb insert. Figure 20(D) showed that all the radioactively labelled single stranded probes migrated to the same position and corresponded to the original 2.5 kb position (Figure 20(C)). The probes made by 3 different ways were used to hybridize different blots. In Figure 21, the probe made by 5' end labelling specifically hybridized to the 2.5 kb but not 3.5 kb and the size markers. In Figure 22, the probe made by 3' end labelling hybridized to the dot blot homogeneously no matter how different the insert sizes were. In Figure 25, the probe made by primer extension with Klenow showed that the probe specifically hybridized to the complementary single strand genome but not the vector DNA.

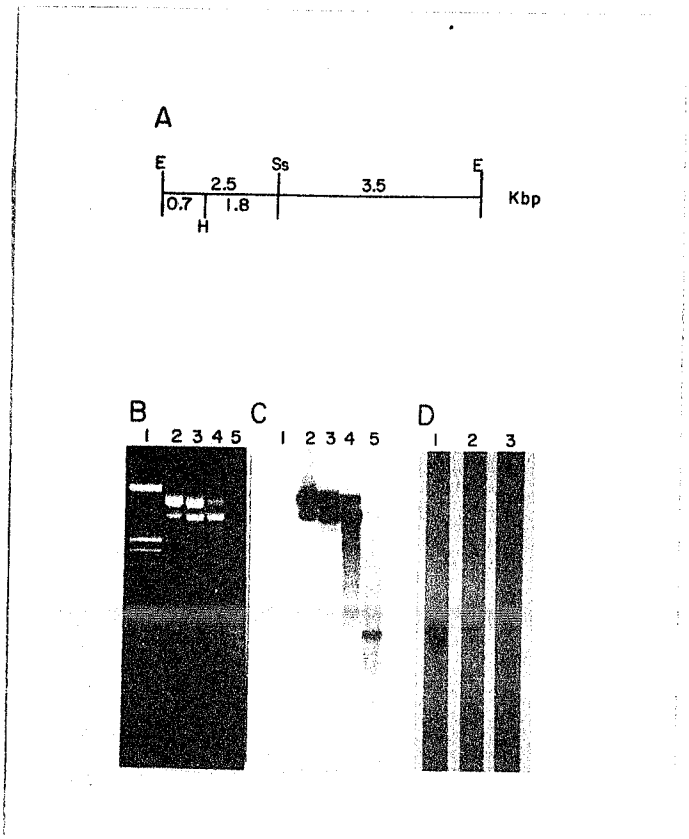

Figure 20. **Making of strand and length specific probes from a 2.5 kb single stranded insert.** Single stranded phage DNA from M13mp18 clone containing a 2.5 kb insert was prepared as described in MATERIALS and METHODS, annealed respectively or together with the synthetic 29 mer and 20 mer which were complementary to each end of the polylinker region of M13mp18 (sequence as described in MATERIALS and METHODS), and then digested with EcoRI and HindIII. 1% low melting point agarose gel was electrophored, blotted, Hybridized with the nick-translated 2.5 kbp fragment as described in MATERIALS and METHODS. Single stranded DNA corresponding to 2.5 kb was isolated and purified from gel. Probes were made by 5' and 3' end labelling or primer extension of single stranded phage as described in MATERIALS and METHODS. (A) Schematic representation of the restriction map and the position of 2.5 kbp fragment. (B) Lane 1, lambda DNA digested with EcoRI and HindIII as size marker; lane 2, single stranded phage DNA; lane 3, single stranded phage DNA annealed with 20 mer and digested with EcoRI; lane 4, single stranded phage DNA annealed with 29 mer and digested with HindIII; lane 5, single stranded phage DNA annealed with 20 mer and 29 mer, and then digested with EcoRI and HindIII. (C) Southern blot and hybridization with the nick-translated 2.5 kbp fragment. (D) Making of strand and length specific probes. Lane 1, single stranded phage DNA was annealed to M13 universal 17 mer, extended with Klenow fragment, and then digested with HindIII. The labelled DNA was denatured before subjected to gel electrophoresis. Lane 2, purified 2.5 kb single stranded DNA was labelled at 5' with T_4 DNA Kinase end as described in MATERIALS and METHODS. Lane 3, purified 2.5 kb single stranded DNA was labelled at 3' end with TdT as described in MATERIALS and METHODS.

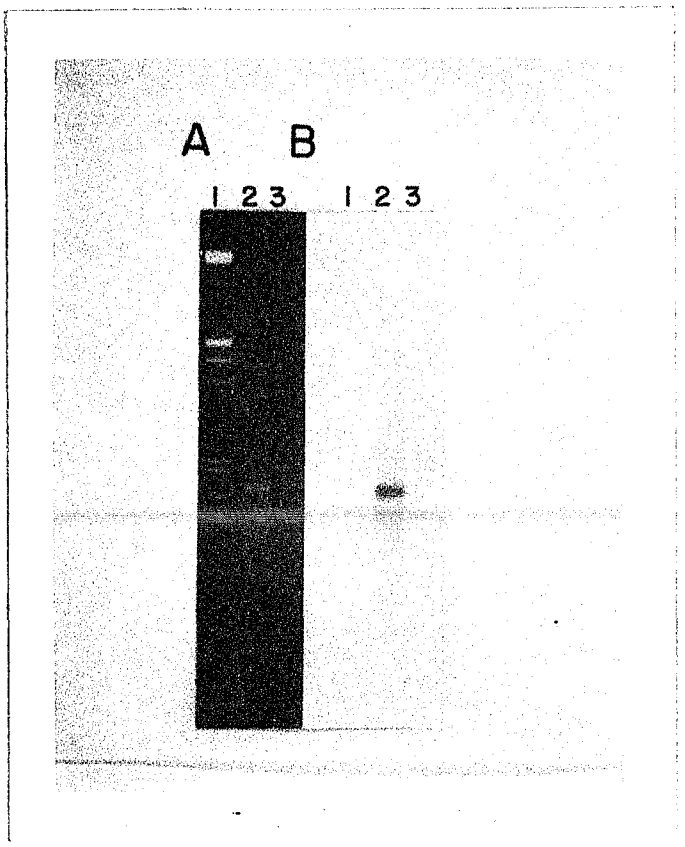

Figure 21. **Southern transfer and hybridization using strand and length specific probes.** Single stranded phage DNAs containing insert with different length were digested with EcoRI and HindIII as described in Figure 20 and subjected to 0.8% agarose gel, blotted, hybridized with a 5' end labelled probe generated as described in Figure 20. (A) Gel electrophoresis of the digests of single stranded phage DNAs. Lane 1, HindIII and EcoRI digests of lambda DNA as size markers; Lane 2, digests of phage DNA with a 2.5 kb insert; Lane 3, digests of phage DNA with a 3.5 kb insert. (B) Southern blot and hybridization with the 5' end labelled probe made from the 2.5 kb DNA.

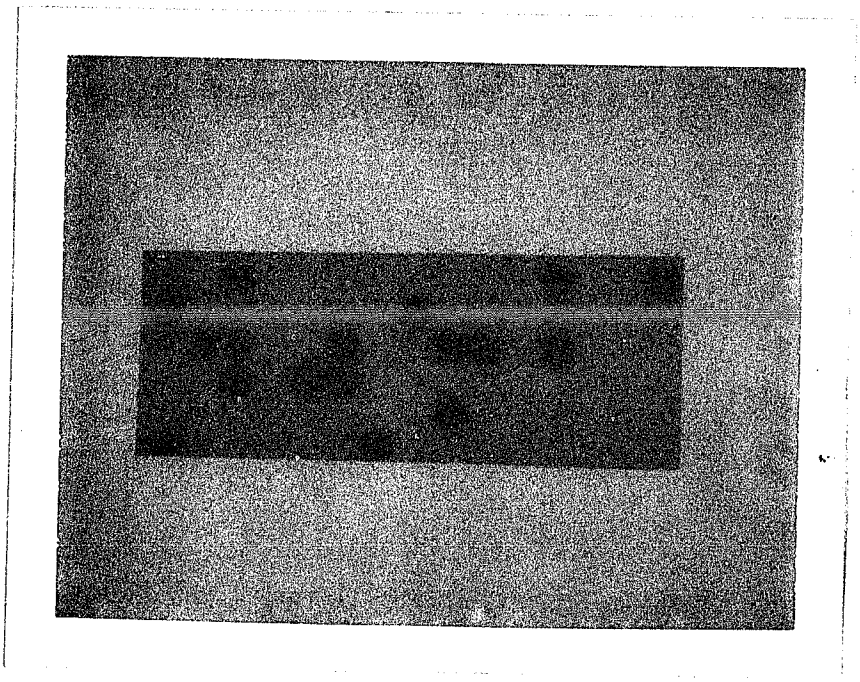

Figure 22. Dot hybridization using strand and length specific probe. A M13 library was generated by shotgun cloning of PstI digest of 6 kbp fragments from rabbit muscle PFK gene. Ninety clones with insert were selected by their phenotypic change. Phages supernatant were spot on nitrocellulose paper and hybridized with the 3' end labelled probe from 2.5 kb single stranded DNA as described in MATERIALS and METHODS and Figure 20.

2. Unidirectional nested DNA sequencing.

Figure 23 show the schematic representation of the improved unidirectional nested sequencing strategy. RF DNA with insert subcloned into A and B sites of polylinker is purified as described in MATERIALS and METHODS. RF DNA is digested with restriction enzyme A, trimed with exonuclease Bal-31 for various time intervals, and digested with restriction enzyme B and X, while vector DNA is prepared by multiply digests with restriction enzymes B, C, and Y. C will create a blunt end, whereas B creates a cohesive end. The purpose to digest with restriction enzymes X and Y is to reduce the possibility of self-ligation of insert to vector and vector by itself. The digests are pooled and an aliquot is used to ligate with digested vector with T_4 DNA ligase and then transformed into JM 107. Phage particles on plates with appropriate numbers are transferred to nitrocellulose paper and hybridized with nick-translated probe. Phage SS or RF DNA with various length of inserts are prepared as described in MATERIALS and METHODS. SS DNA or RF DNA are sequenced using M13 universal sequencing primer or reverse sequencing primer as described in MATERIALS and METHODS. A M13 library created by using this improved strategy was screened by in situ hybridization. The results were shown in Figure 24. Fifty positive ones were shown on the dish having a total 55 phage particles including 50 white ones, while 100 positive ones were shown on the dish having a total 105 phage particles including 100 white ones. This result showed that multiple digestion with restriction

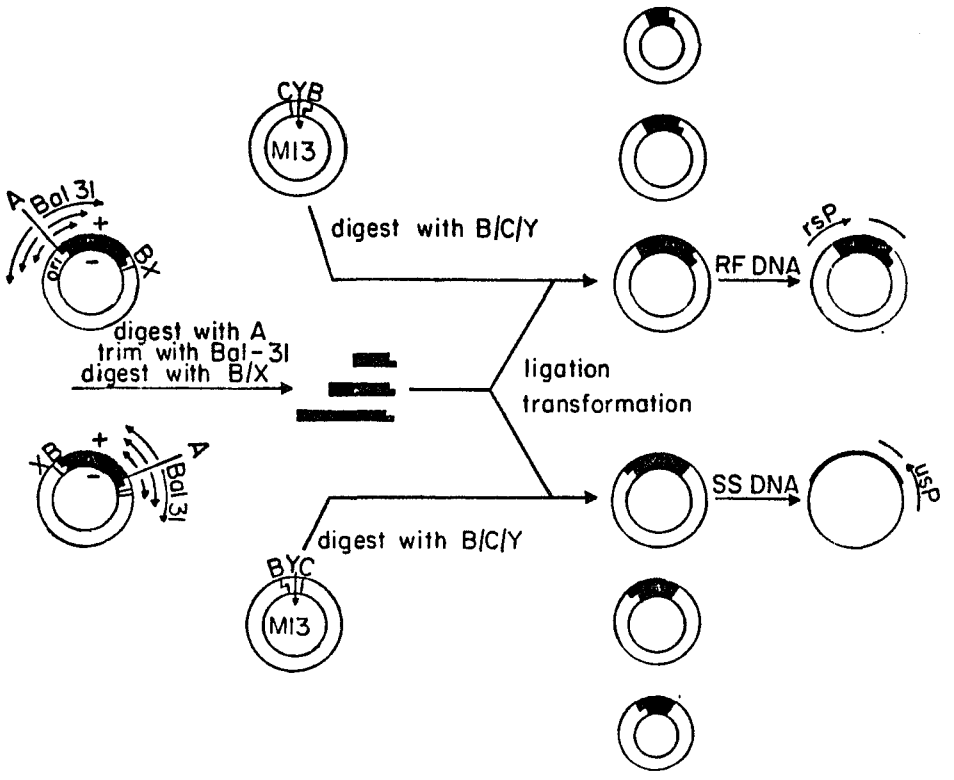

Figure 23. Schematic representation of an improved unidirectional nested DNA sequencing strategy.

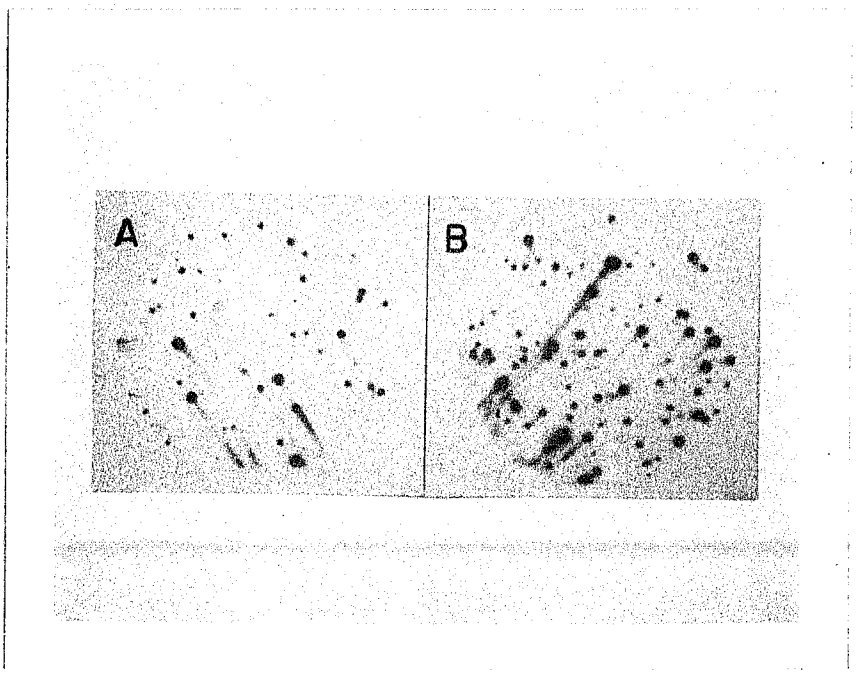

Figure 24. In situ hybridization of M13 clones containing the Bal-31 treated fragments of rabbit muscle PFK gene. A 2.5 kbp, EcoRI/SstI DNA fragment purified from the pBR325/6.0 kbp clone was digested with PstI, subcloned into the polylinker of M13mp18. Phage particles with insert were selected by their lac phenotypic change. RF DNA purified as described in MATERIALS and METHODS was digested with EcoRI, trimmed with exonuclease Bal-31 and subcloned into the SstI/HincII sites of M13mp18 as described in RESULTS. After transformation, fresh phage particles on petri dishes with the numbers from 50 to 100 were selected, blotted, hybridized with the nick-translated 2.5 kbp probe as described in MATERIALS and METHODS. (A) A dish with a total 55 phage particles and 50 white ones. (B) A dish with a total 105 phage particles and 100 white ones.

enzyme resulted a low background in transformation. Eighteen positive clones were selected randomly lysed by SDS and subjected to gel electrophoresis. The results on Figure 25(A) showed that all the inserts ranged within 2.5 kbp in size. When the blot was hybridized with the strand and length specific probe, it further demonstrated that all of the inserts showed the same polarity and came from the same 2.5 kbp fragment(Figure 25(B)). Template DNAs from these 18 clones were purified and sequenced. DNA sequence was analyzed by computer. Results was shown in Figure 12(C). It showed that DNA sequences were well overlapped and continuous.

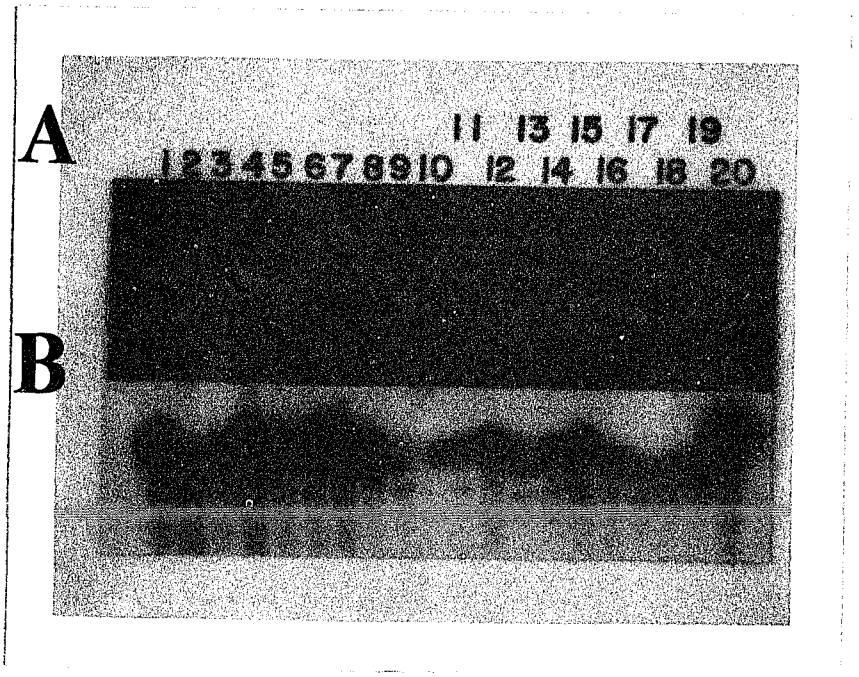

Figure 25. Southern transfer and hybridization of single atranded M13 clones generated by Bal-31 using strand and length specific probe. (A) Eighteen M13 clones (lane 2 to 19) generated as described in Figure 24 were selected, lysed, subjected to 0.4% agarose gel electrophoresed and then stained by ethidium bromide as described in MATERIALS and METHODS. Single stranded M13mp18/2.5 kbp and M13mp18 were mixed, lysed, and run on lane 1 and 20 as size markers. (B) The phage DNAs were transferred to nitrocellulose paper, hybridized with the probe generated by primer extension of single stranded M13mp18/2.5 kbp as described in RESULTS.

DISCUSSION

I. RMPFK gene --- DNA sequence, genomic organization and gene evolution

1. Is rabbit muscle PFK evolved from a prokaryotic progenitor by gene duplication and divergence?

Mammalian PFKs have been speculated to have evolved from a prokaryotic progenitor gene by gene duplication and divergence (Poorman et al., 1984). This speculation is based on the comparisons of RMPFK and BsPFK. If we accept this speculation, two criteria should be met: (1) the PFK gene is composed of two duplicated halves and (2) these two halves are not only similar each other, but also similar to the BsPFK gene. To evaluate the first criterion, we first focus on the relation between the exons of the gene and the functional domains of this protein. Binding sites in rabbit muscle PFK are classified into two types: the catalytic sites including sites binding substrate F6P and substrate ATP, and effector sites including those for the binding of activators ADP and FBP and inhibitor ATP (Poorman et al., 1984, Hellinga and Evans, 1985). We have grouped these sites as sugar binding sites including sites binding F6P and FBP and nucleotide binding sites such as ADP activation sites and ATP inhibition sites. Sites for binding substrate ATP are grouped separately for the sake of clarity. When the amino acid residues of the enzyme are labelled with these designations (Figure 3, boxed residues), it becomes noticeable that these sites form subdomain, each enclosed within an exon. With few exceptions, almost no

introns are found to interrupt these subdomains. More strikingly, when the exons in the map (Fig. 8) are identified by their encoded subdomains a clear pattern of gene duplication is revealed. The organization of exons in the N-half of the gene is duplicated in the C-half. Exon 13 corresponds to a fusion of exons 1 and 2 while exon 17 represents a fusion of exons 6 and 7. Exon 4 contributes only substrate ATP binding sites while exon 15 contains the corresponding diverged residues. The order of exon 21 and 22 appears to be an inversion of exons 11 and 12 or vice versa. This inversion is possibly due to the loss of an intron between Lys-324 and Ile-325 in C-half (Figure 3). Based on the duplication pattern, the number of exons in N-half is the same as that of the C-half (Table 9). Although the length of the introns separating the exons are quite different in the two halves (Table 2), the size of these corresponding exon pairs is indeed similar in this gene (Table 9). The homology (43%) of the coding sequence between these two halves is even higher than the homology (30%) of protein sequence (Table 10). Futhermore, two halves show a similar predominant codon usage (Table 8). The evidence discussed so far fit the first criterion that the RMPFK gene is composed of two duplicated halves.

At the protein level, it has been shown that the N-half of RMPFK is more similar to BsPFK as compared to the C-half (Poorman et al., 1984; also see Table 10). However, by DNA sequence comparison (Table 10), our results suggest that the C-half of RMPFK is more similar to BsPFK. Nevertheless, both N- and C-halves of RMPFK have a significant homology to that of BsPFK in the DNA

Table 9. Comparisons of size and identity between corresponding exons in the two halves of the rabbit muscle PFK gene.

Exon	Residue number encoded	exon size bp	% identity Amino acids	Nucleotides
1+2	52	156	54	55
13	50	150		
3	26	78	33	38
14	24	71		
4	63	190	31	40
15	29	88		
5	55	166	28	46
16	51	153		
6+7	51	154	40	49
17	55	165		
8	32	96	5	18
18	21	62		
9	31	93	40	52
19	37	112		
10	42	126	29	50
20	33	100		
11	22	65	18	37
22	47	142		

Table 10. The DNA and amino acid homology comparison of related PFKs. Amino acid homology is shown by italic number and DNA homology is shown by regular number.

	EcPFK	BsPFK	RMPFK(N)	RMPFK(C)
EcPFK		*53*	*39*	*30*
BsPFK	57		*42*	*33*
RMPFK(N)	47	43		*30*
RMPFK(C)	43	48	43	

level. The evidence from the comparisons of DNA sequences satisfy the second criterion that the two halves of RMPFK gene are not only similar each other, but also similar to BsPFK gene. The data so far discussed support the hypothesis that mammalian PFK gene evolved from a prokaryotic progenitor and divergence.

2. Intron/exon organization of RMPFK gene.

Gilbert (Gilbert, 1978) proposed that introns divide genes into exons which might correspond to structural or functional domains in the final proteins. Recombination within introns could assort these functions independently. Individual exons could be thought of as building blocks; a single block might be duplicated and used several times in one or more genes. Our results show that most introns of RMPFK gene are localized between or at the ends of the secondary structural elements (Figure 19) and each exon represents a functional subdomain (Figure 12(A)). There are two extreme views for the origin of intron. Doolittle (Doolittle, 1978) proposed that introns were in the preexisting gene. This hypothesis is based on the assumptions; (1) a splicing apparatus existed in the first cells and (2) introns were lost from prokaryotes as their genomes became streamlined under the selective pressure of rapid DNA replication, while introns were retained in eukaryotes. Contrastly, Orgel and Crick (Orgel and Crick, 1980) proposed that introns were laterly added at some point to break up preexisting genes. The latter hypothesis is on the assumption that prokaryotic genes resemble the preexisting genes and that the addition of introns and the

occurrence of splicing apparatus emerged later during the evolution of eukaryotic genes. However, the demonstration of RNA self-splicing events in tetrahemena by Cech (Cech, 1986) and the finding of similar exon/intron motifs in triose phosphate isomerases from different sources by Gilbert and his coworkers (Gilbert et al., 1986) have made the first hypothesis more attractive and acceptable. If we accept Doolittle's hypothesis, we should see similar exon/intron motifs for the two halves of RMPFK gene. However, our results are contrary to this prediction. Introns on the C-half of this gene are not located at identical positions as those in the N-half (Figure 3). This result suggests that duplicated two halves may not have identical exon/intron motif, because intron position may be changed due to intron deletion or recombination (Gilbert et al., 1986). Furthermore, mutation at splicing junctions would result splicing to use other cryptic splicing junctions; these junctions are inactive before mutation (Ruskin et al., 1984). This migration and change of intron position result part of intron sequences being exon sequences after splicing. This added exon sequences would then code extra amino acid residues in the final protein product. There are four regions in the rabbit muscle PFK sequence which stand out as extra residues (Figure 3) when the two halves of RMPFK are aligned with two bacterial PFK sequences for maximum homology. All these four extra regions have intron to interrupt the sequences. These results agree with our suggestion that these extra amino acid residues evolved from intron sequences and the migration of splicing junctions over evolution time.

3. Tissue-specific codon usage in RMPFK gene.

Newgard and his coworkers (Newgard et al., 1986) reported that cDNAs of glycogen phosphorylase from liver and muscle exhibited a remarkable divergence in G+C content at the third codon positions (86% to 60%). They further surveyed other mammalian cDNAs and found that this bias in codon usuage patterns in liver and muscle was in general. This finding leads to the suggestion that codon usage may play an important role in specific gene expression of tissues. RMPFK gene has 75% of codons have G or C at the third codon position (Table 8). These results support that this gene so far discussed are the one encoding muscle PFK rather than other isozymes.

4. Full-length cDNA and isozyme gene(s): from known to unknown.

The catalytic and allosteric sites of RMPFK so far referred are deduced from the crystallographic structure of BsPFK (Poorman et al., 1984). To construct a full-length cDNA to verify those putative sites using site-directed mutagenesis approach seems to be an unavoidable way. However, several attemps to make the full-length PFK cDNA failed due to unknown reasons (Vora et al., 1986; French, personal communication; Valdez, personal communication). In addition to the sequences of all exons, we also report the restriction sites and possible secondary structure in this coding sequences (Figure 26). This coding sequence in turn represents the mRNA of rabbit muscle PFK gene and this information will be very helpful to construct a full-length cDNA. Furthermore, DNA fragments of RMPFK gene have been subcloned into

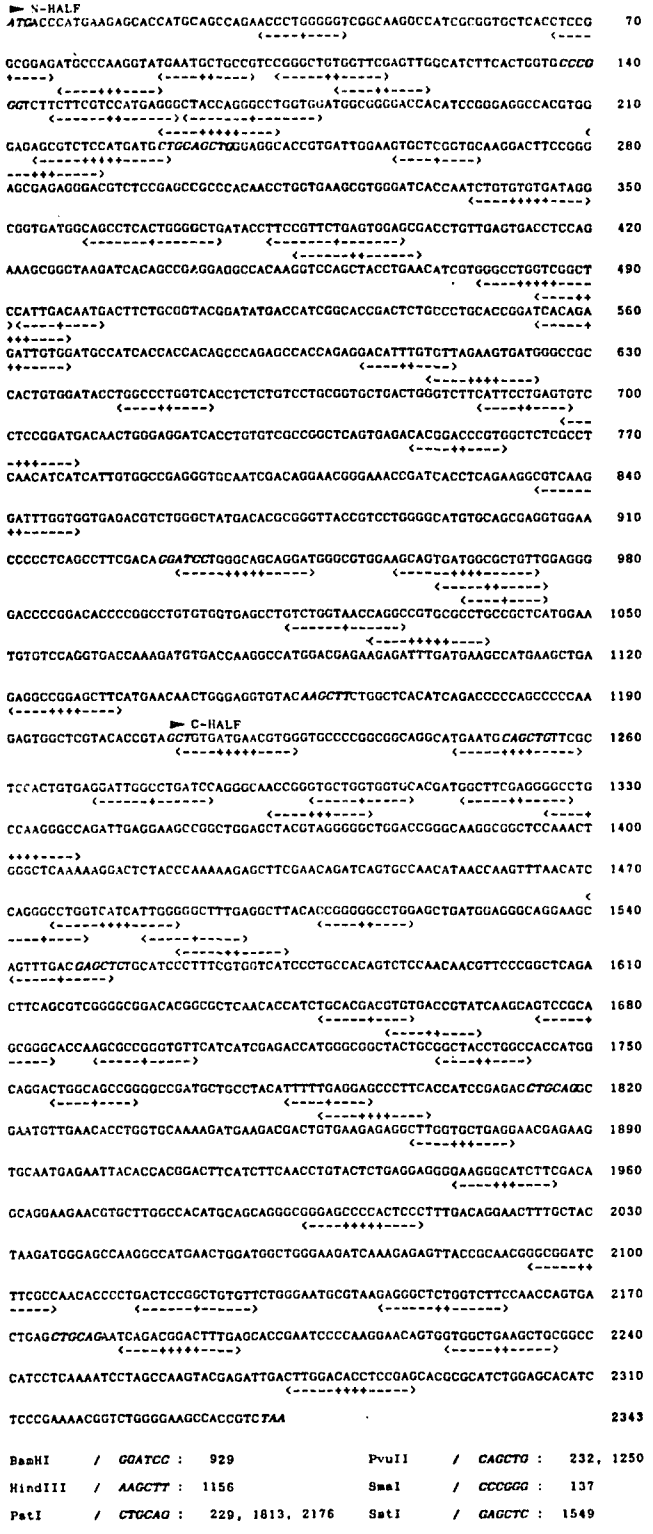

Figure 26. The possible secondary structural position in the coding sequence of rabbit muscle PFK gene. The numbers in the right margin correspond to the sequence from the first codon ATG (number with 1) to the termination codon TAA (number with 2343). The solid arrows indicate the first codon for N- and C-halves of the protein. Sequences of restriction sites are indicated by italic and listed on the bottom. Secondary structures were computed by using Staden program (Staden, 1980). The loop size was set from 0 to 5, stem size was set at 5. The angled brackets spaned the region of possible secondary structure. "+" indicates the loop sequence and region.

M13 and pBR vectors. These gene fragments will serve as invaluable probes for other PFK isozyme genes (or pseudogenes).

5. Rabbit muscle PFK: from DNA to protein.

The amino acid sequence of the rabbit muscle PFK is revised and completed from the DNA sequence (Figure 8). The first ATG codon triplet codes for the initiation methionine residue which is immediately followed by the threonine residue. This NH_2-terminal methionine residue will be cleaved from the nascent polypeptide by an aminopeptidase (Yoshida & Lin, 1972). Our DNA sequence shows that there is no leader peptide in this protein. Moreover, no matter with or without the initiation methionine residue, matured RMPFK still retain a stabilizing residue in the first NH_2-terminal position according to the N-end rule (Bachmair et al., 1986). The long-lived property predicted from the N-end rule is consistent with the stable property as observed (Uyeda, 1979; Foe & Kemp, 1986). High percentages of non-charged amino acid residues (Ala, Gly, Leu, Thr, Val) (Table 7) reflect rabbit muscle PFK has a pI close to pH 8.0 (Uyeda, 1978; Kemp, 1986). This pI explains why rabbit muscle PFK is stable at pH 8.0 but unstable at lower pH (Uyeda, 1979; Foe & Kemp, 1986).

II. Techniques developed in this work

1. Rapid mapping strategy and applications.

DNA fragment inserted into a polylinker of phage or plasmid vector can be rapidly mapped for their restriction sites. In the beginning of this project, the 17 kbp rabbit muscle PFK gene were

mapped by 13 restriction endonucleases using this new strategy. Our DNA sequence has proved the correctness of this map. It is evident that this method is simple and direct. No steps for purification of DNA fragment from gel are needed. Another advantage is that once the restriction map of the insert is worked out, DNA sequencing can be carried on by Maxam-Gilbert (Maxam & Gilbert, 1977) or dideoxy sequencing (Sanger et al., 1977) method without further subcloning. However, this method also suffers the limitation: only enzymes which have the restriction sites in the polylinker can be used, except a few cases (ex. PvuII). Otherwise, the fragments coming from the vector part will interfere the interpretation of autoradiogram.

2. A technique to make strand and length specific probes and its applications.

We have developed a technique for making the strand and length specific probes. These probes have been successfully used to screen the M13 libraries and proved to be very specific. This technique can be applied to isolate full-length single stranded DNA for S1 mapping, viroid study or site-directed mutagenesis.

3. The advantages of the improved nonrandom DNA sequencing strategy.

This strategy has been proved to possess the following advantages:

(1) It is not necessary to isolate Bal-31 trimmed fragments from gel.

(2) Multiple digestions greatly decrease the possibility of

self-ligation of the insert with original vector or the new vector by itself.

(3) If M13 phage loses either the origin of replication or the gene II sequences, it won't be viable. This characteristic provides a good automatic screening system. Each viable white phage particle theoretically represents a positive clone.

(4) DNA sequence will be read from a predefined end and strand. This will dramatically decrease time and difficulties from computer analysis.

(5) Both single stranded and double stranded template can be used for bidirectional sequencing.

LITERATURE CITED

Bachmair, A., Finley, D., Varshavsky, A. 1986. In vivo Half-Life of A Protein Is A Function of Its Amino-Terminal Residue. Science 234: 179-186.

Barnes, W. M. 1977. Plasmid Detection and Sizing in Single Colony Lysate. Science 195: 393-394.

Bauer, B. A., Younathan, E. S. 1984. Decreased Phosphofructokinase Activity in Skeletal Muscle of Diabetic Rats. Clinical Physiology and Biochemistry 2: 137-145.

Biggin, M.D., Gibson, T. J., and Hong, G. F. 1983. Buffer Gradient Gel and ^{35}S Label as an Aid to Rapid DNA Sequence Determination. Proc. Natl. Acad. Sci. USA 80: 3963-3965.

Birnboim, H. C., Doly, J. 1979. A Rapid Alkaline Extraction Procedure for Screening Recombinant Plasmid DNA. Nucl. Acids Res. 7: 1513-1523.

Blake, C. C. F. 1983. Exons-Present from the Beginning? Nature 306: 535-537.

Blattner, F. R., Williams, B. G., Blechl, A. E., Denniston-Thompson, K. Faber, H. E., Furlong, L.-A. Grunwald, D. J., Kiefer, D. O., Moore, D. D., Sheldon, E. L., Smithies, O. 1977. Charon Phages: Safer Derivatives of Bacteriophage Lambda for DNA Cloning. Science 196: 161

Bloxham, D. P., Lardy, H. A. 1973. Phosphofructokinase. The Enzymes pp. 239-278. edited by Boyer, P. D., Academic press.

Bolivar, F. 1978. Construction and Characterization of New Cloning Vehicles III Derivatives of Plasmid pBR322 Carrying Unique EcoRI Sites for Selection of EcoRI Generated Recombinant DNA Molecules. Gene 4: 121-136.

Branden, C. I., Eklund, H., Cambillan, C., and Pryor, A. J. 1984. Correlations of Exons with Structure Domains in Alcohol Dehydrogenase. EMBO J. 3: 1307-1310.

Breathnach, R., Chambon, P. 1981. Organization and Expression of Eucaryotic Split Genes Coding for Proteins. Ann. Rev. Biochem. 50: 349-383.

Carroll, S. L., Bergsma, D. J., Schwartz, R. J. 1986. Structure and Complete Nucleotide Sequence of the Chicken alpha-Smooth Muscle (Aortic) Actin Gene. J. Biol. Chem. 261: 8965-8976.

Chaconas, G. O., van de Sande, J. H. 1980. 5'-^{32}P Labelling of RNA

and DNA Restriction Fragments. Methods in Enzymology. 65: 75-85.
Chang, S. H., French, B. A., Lee, C. P., Younathan, E. S., and Putney, S. D. 1984. Molecular Cloning and Characterization of Rabbit Phosphofructokinase Genes. Fed. Proc. Abstract No. 1791, St. Louis.

Covarrubias, L., Cervantes, L., Covarrubias, A., Soberon, X., Vichide, I., Blanco, A., Bolivar, F. 1981. Construction and Characterization of New Cloning Vehicles V. Characterization and Coding Properties of pBR322 and Several Deletion Derivatrives Including pBR327 and pBR328. Gene 14: 25-35.

Dale, R. M. K., McClure, B. A., Houchins, J. P. 1985. A Rapid Single-Stranded Cloning Strategy for Producing a Sequential Series of Overlapping Clones for Use in DNA Sequencing: Application to Sequencing the Corn Mitochondrial 18S rDNA. Plasmid 13: 31-40.

Darnell, J. E., and Doolittle, W. F., 1986. Speculations on the Early Course of Evolution. Proc. Natl. Acad. Sci. USA 83: 1271-1275.

Davidson, M., Collins, M., Byrne, J., Vora, S. 1983. Alternation in Phosphofructokinase Isozymes during Early Human Development. Biochem. J. 214: 703-710.

Davidson, M., Miranda, A. F., Bender, A. N., DiMauro, S., Vora, S. 1983. Muscle Phosphofructokinase Deficiency. J. Clin. Invest 72: 545-550.

Deininger, P. L. 1983. Random Subcloning of Sonicated DNA Sequence Analysis. Anal. Biochem. 129: 216-223.

Evans, P. R., Farrants, G. W., and Hudson, P. J. 1981. Phosphofructokinase: Structure and Control. Phil. Trans. R. Soc. Lond. B 293: 53-62.

Evans, P. R., and Hudson, P. J. 1979. Structure and Control of Phosphofructokinase from B. stearothermophilus. Nature 279: 500-504.

Evans, P. R., Farrants, G. W., Lawrence, M. C. 1986. Crystallographic Structure of Allosterically Inhibited Phosphofructokinase at 7 Å Resolution. J. Mol. Biol. 191: 713-720.

Fitch, W. M., Smith, T. F., Ralph, W. W. 1983. Mapping the Order of DNA Restriction Fragments. Gene 22: 19-29.

Foe, L. G., Kemp, R. G. 1984. Isozyme Composition and Phosphorylation of Brain Phosphofructokinase. Arch. Biochem. Biophys. 228: 503-511.

Foe, L. G., Kemp, R. G. 1986. Isolation and Characterization of Phosphofrutokinase C from Rabbit Brain. (in press)

Fothergill-Gilmore, L. A. 1986. The Evolution of the Glycolytic Pathway. TIBS 11: 47-50.

Fraenkel, D. G. 1986. Mutants in Glucose Metabolism. Ann. Rev.

Biochem. 55: 317-337.

French, B. A., and Chang, S. H. Molecular Cloning and Sequencing of the Gene for Phosphofructokinase from B. stearothemophilus. (in preparation).

Garland, P. B., Randle, P. J., Newsholme, E. A. 1963. Citrate as an Intermediary in the Inhibition of Phosphofructokinase in Rat Heart Muscle by Fatty Acids, Ketone Bodies, Pyrute, Diabetes and Starvation. Nature 200: 169-170.

Gatti, R. A., Concannon, P., Salser, W. 1984. Multiple Use of Southern Blots. BioTechniques. May/June.

Gilbert, W. 1978. Why Genes in Pieces? Nature 271: 501.

Gilbert, W., Marchionni, M., and McKnight, G. 1986. On the Antiquity of Introns. Cell 46: 151-154.

Go, M. 1981. Correction of DNA Exonic Regions with Protein Structure Units in Haemoglobin. Nature 291: 90-92.

Hanahan, D. 1983. Studies on Transformation of Escherichia coli with Plasmids. J. Mol. Biol. 166: 557-580.

Hellinga, H. W., and Evans, P. R. 1985. Nucleotide Sequence and High Level Expression of Major E. coli Phosphofructokinase. Eur. J. Biochem. 149: 363-373.

Hers, H. G., Hue, L., and Schaftingen, E. V. 1982. Fructose 2,6-bisphosphate. Trends in Biochem. Sci. 7: 329-331.

Hofmann, E. 1978. Phosphofructokinase --- A Favorite of Enzymologists and of Students of Metabolic Regulation. Trends Biochem. Sci. 8: 145-147.

Hohn, B., Murray, K. 1977. Packaging Recombinant DNA Molecules into Bacteriophage Particles in vitro. Proc. Natl. Acid Sci. USA 74: 3259-3263.

Ibsen, K. H., Fishman, W. H. 1979. Developmental Gene Expression in Cancer. Biochimica et Biophysica Acta. 560: 243-280.

Keller, E. B., Noon, W. A. 1984. Intron Splicing: A Conserved Internal Signal in Introns of Animal Pre-mRNAs. Proc. Natl. Acad. Sci. USA 81: 7417-7420.

Kemp, R. G., Krebs, E. G. 1967. Binding of Metabolites by Phosphofructokinase. Biochemistry 6: 423-434.

Kemp, R. G., and Foe, L. G. 1983. Allosteric Regulatory Properties of Muscle Phosphofructokinase. Mol. Cell. Biochem. 57: 147-154.

Koerner, T. A. W., Jr., Voll, R. J., and Younathan, E. S. 1977. A Proposed Model for the Regulation of Phosphofructokinase and Fructose 1,6-Bisphosphatase Based on Their Reciprocal Anomeric Specificity. FEBS Letters 84: 207-213.

Kolb, E., Hudson, P. J., and Harris, J. I. 1980. Phosphofructokinase: Complete Amino Acid Sequence of the Enzyme from B. stearothermophilus. Eur. J. Biochem. 108: 587-597.

Kyte, J., Doolittle, R. F. 1982. A Simple Method for Displaying the Hydropathic Character of a Protein. J. Mol. Biol. 157: 105-132.

Lagrimini, L. M., Brentano, S. T., and Donelson, J. E. 1984. A DNA Sequence Analysis Package for the IBM Personal Computer. Nucleic Acids Res. 12: 605-614.

Lau, P. P., Gray, Jr. H. B. 1979. Extracellular Nucleases of Alteromonas espejina Bal 31. IV: The Single Strand Specific Deoxyriboendonuclease Activity as A Probe for Regions of Altered Secondary Structure in Negatively and Positively Supercoiled Closed Circular DNA. Nucl. Acids Res. 6: 331

Levitt, M., Chothia, C. 1976. Structural Patterns in Globular Proteins. Nature 261: 552-558.

Lonberg, N., and Gilbert, W. 1985. Intron/Exon Structure of the Chicken Pyruvate Kinase Gene. Cell 40: 81-90.

Maniatis, T., Fritsch, E. F., and Sambrook, J. 1982 Molecular Cloning, (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K., and Efstratiadis, A. 1978. The Isolation of Structure Genes from Libraries of Eucaryotic DNA. Cell 15: 687-701.

Marchionni, M., and Gilbert, W. 1986. The Triosephosphate Isomerase Gene from Maize: Inrons Antedate the Plant-Animal Divergence. Cell 46: 133-141.

Matteucci, M. D., and Caruthers, M. H. 1981. Synthesis of Deoxyoligonucleotides on a Polymer Support. J. Am. Chem. Soc. 103: 3185-3191.

Mattice, W. L., Srinivasan, G., Santiago, G. 1980. Helix End Effects in Block Copolypeptides, Proteins, and Protein-Detergent Complexes. Macromolecules 13: 1254-1260.

Maxam, A. M., and Gilbert, W. 1977. A New Method of Sequencing DNA. Proc. Natl. Acad. Sci. USA. 74: 560-564.
Meinkoth, J., Hahl, G. 1984. Hybridization of Nucleic Acids Immobilized on Solid Supports. Anal. Biochem. 138: 267-280.

Messing, J. 1983. New M13 Vectors for Cloning. Methods in Enzymo-

logy 101: 20-78.

Michelson, A. M., Blake, C. C. F., Evans, S. T., and Orkin, S. H. 1985. Structure of the Human Phosphoglycerate Kinase Gene and the Nucleotide Binding Domain. Proc. Natl. Acad. Sci. USA 82: 6965-6969.

Misra, T. K. 1985. A New Strategy to Create Ordered Deletions for Rapid Nucleotide Sequencing. Gene 34: 263-268.

Mount, S. M. 1982. A Catalogue of Splice Junction Sequence. Nucl. Acids Res. 10: 459-472.

Norrander, J., Kempe, T., Messing, J. 1983. Construction of Improved M13 Vectors Using Oligodeoxynucleotide-directed Mutagenesis. Gene 26: 101-106.

Nussinov, R. 1981. Eukaryotic Dinucleotide Preference and Theier Implication for Degenerate Codon Usage. J. Mol. Biol. 149: 125-131.

Park, I., Schaeffer, E., Sidoli, A., Baralle, F., Cohen, G. N., Zakin, M. 1985. Organization of the Human Transferrin Gene: Direct Evidence That It Orignated by Gene Duplication. Proc. Natl. Acad. Sci. USA 82: 3149-3153.

Parmeggiani, A., Luft, J. H., Love, D. S., Krebs, E. G. 1966. Crystallization and Properties of Rabbit Skeletal Muscle Phosphofructokinase. J. Biol. Chem. 241: 4625-4637.

Passonneau, J. V., Lowry, O. H. 1962. Phosphofrucotokinase and the Pasteur Effect. Biochem. Biophys. Res. Commun. 7: 10-15.

Poncz, M., Solowiejczyk, D., Ballantin, M., Schwartz, E., Surrey, S. 1982. "Nonrandom" DNA Sequence Analysis in Bacteriophage M13 by the Dideoxy Chain-termination Method. Proc. Natl. Acad. Sci. USA 79: 4298-4302.

Poorman, R. A., Randolph, A., Kemp, R. G., and Heinrikson, R. L. 1984. Evolution of Phosphofructokinase-Gene Duplication and Creation of New Effector Sites. Nature 309: 467-469.

Putney, S., P. D., Herlihy, W. C., and Schimmel, P. 1983. A New Troponin T and cDNA Clones for 13 Different Muscle Proteins Found by Shotgun Sequencing. Nature 302: 718-721.

Putney, S., Herlihy, W., Royal, N., Pang, H., Aposhian, H. V., Pickering, L., Relagaje, B., Biemann, K., Page, D., Kuby, S., and Schimmel, P. 1984. Rabbit Muscle Creatine Phosphokinase. J. Biol. Chem. 259: 14317-14320.

Rackwitz, H.-R., Zehetner, G., Frischauf, A.-M., Lehrach, H. 1984. Rapid Restriction Mapping of DNA Cloned in Lambda Phage Vectors. Gene 30: 195-200.

Reed, R., Maniatis, T. 1986. A Role for Exon Sequences and Splice-Site Proximity in Splice-Site Selection. Cell 46: 681-690.

Rigby, P. W. J., Dieckmann, M., Rhodes, C., Berg, P. 1977. Labelling Deoxyribonucleic Acid to High Specific Activity in vitro by Nick Translation with DNA Polymerase I. J. Mol. Biol. 113: 237-251.

Rogers, J. 1985. Exon Shuffling and Intron Insertion in Serine Protease Genes. Nature 315: 458-459.

Ruskin, B., Krainer, A. R., Maniatis, T., and Green, M. R. 1984. Excision of An Intact Intron as A Novel Lariat Structure during Pre-mRNA Splicing in vitro. Cell 38: 317-331.

Sanger, F., Coulson, A. R., Barrel, B. G., Smith A. J. H., Roe, B. A. 1980. Cloning in Single-Stranded Bacteriophage as An Aid to Rapid DNA Sequencing. J. Mol. Biol. 143: 161-178.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA Sequencing with Chain-Terminating Inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Schimid, C. W., Jelinek, W. R. 1982. The Alu Family of Dispersed Repetitive Sequences. Science 216: 1065-1070.

Sharp, P. A. 1983. Conversion of RNA to DNA in Mammals: Alu-like Elements and Pseudogenes. Nature 301: 471-472.

Singer, M. F. 1982. SINEs and LINEs: Highly Repeated Short and Long Interspersed Sequences in Mammalian Genomes. Cell 28: 433-434.

Smith, H. O., Birnstiel, M. L. 1976. A Simple Methods for DNA Restriction Site Mapping. Nucl. Acids Res. 3: 2387-2400.

Smith, M. Synthetic Oligodeoxyribonucleotides as Probes for Nucleic Acids and as Primers in Sequence Determination. (Methods of RNA and DNA Sequencing, Weissmann, S. M., Praegar Scientific Press New York, N.Y.)

Soberon, X., Covarrubias, L., Bolivar, F. 1980. Construction and Characterization of New Cloning Vehicles, IV. Deletion Derivatives of pBR322 and pBR325. Gene 9: 287-305.

Southern, E. 1975. Detection of Specific Sequences among DNA Fragments Seperated by Gel Electrophoresis. J. Mol. Biol. 98: 503-517.

Staden, R. 1980. A New Computer Method for the Storage and Manipulation of DNA Gel Reading Data. Nucl. Acids Res. 8: 3673-3694.

Staden, R. 1986. The Current Status and Portability of Our Sequence Handling Software. Nucl. Acids Res. 14: 217-231.

Stone, E. M., Rothblum, K. N., and Schwartz, R. J. 1985. Intron-Dependent Evolution of Chicken Glyceradehyde Phosphate Dehydrogenase

Gene. Nature 313: 498-500.

Stryer, L. 1981. "Biochemistry" pp.266, W. H. Freeman and Company Press.

Sun, L., Paulson, K. E., Schmid, C. W., Kadyk, L., Leinwand, L. 1984. Non-Alu Family Interspersed Repeats in Human DNA and Their Transcriptional Activity. Nucl. Acids Res. 12: 2669-2690.

Thrasher, J. R., Cooper, M. D., Dunaway, G. A. 1981. Developmemtal Changes in Heart and Muscle Phosphofructokinase Isozymes. J. Biol. Chem. 256: 7844-7848.

Tsai, M. Y., Kemp, R. G. 1973. Isozymes of Rabbit Phosphofructokinase. J. Biol. Chem. 248: 785-792.

Uyeda, K. 1979. Phosphofructokinase. Adv. Enzym. Related Area Mol. Biol. 48: 193-244.

Uyeda, K. 1973. Rabbit Muscle Phosphofructokinase: the Kinetic Mechanism of Action and the Equilibrium Constant. J. Biol. Chem. 248: 7852-7859.

Vieria, J., Messing, J. 1982. The pUC Plasmids, An M13mp7-derived System for Insertion Mutagenesis and Sequencing with Synthetic Universal Primers. Gene 19: 259-268.

Vora, S., Hong, F., and Olender, E. 1986. Isolation of A cDNA for Human Muscle 6-Phosphofructokinase. Biochem. Biophys. Res. Commun. 135: 615-621.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 Phage Cloning Vector and Host Strains: Nucleotide Sequences of the M13mp18 and pUC19 Vectors. Gene 33: 103-119.

APPENDIX 1. Preparation and composition of reagents

Solution (General)	Preparation /Composition
1 M Tris	Dissolve 121.1 g Tris base in 800 ml of H_2O. Adjust pH to the desired value by adding concentrated HCl. Adjust final volume to 1 l, sterilize by autoclaving.
0.5 M EDTA (pH 8.0)	Add 186.1g of disodium ethylene diamine tetraacetate.$2H_2O$ to 800 ml of H_2O. Stir vigorously on a magnetic stirrer. Adjust pH to 8.0 with NaOH. Adjust final volume to 1 l. Sterilize by autoclaving.
1 M $MgCl_2$	Dissolve 203.3 g of $MgCl_2.6H_2O$ in 800 ml of H_2O. Adjust volume to 1 liter. sterilize by autoclaving.
1 M $MnCl_2$	19.791 g/100 ml in H_2O, sterilize by filtration.
3 M NaOAC (pH 5.2)	Dissolve 408.1 g of sodium acetate.$3H_2O$ in 800 ml of H_2O. Adjust pH to 5.2 with glacial acetic acid. Adjust of volume to 1 liter. Sterlize by autoclaving.
7.5 M NH_4OAC	57.81 g/100 ml in H_2O. Sterilize by filtration.
STE	0.1 M NaCl, 10 mM Tris-HCl (pH 7.8), 1 mM EDTA
Solution I	50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA, 5 mg/ml lysozyme (add lysozyme powder right before use).
Solution II	0.2 N NaOH, 1% SDS (use and made freshly).
5 M KOAC (pH 4.8) (Solution III)	To 60 ml of 5 M potassium acetate add 11.5 ml of glacial acetic acid and 28.5 ml of H_2O. The resulting solution is 3 M with respect to potassium and 5 M with respect to acetate.
20xSSC	Dissolve 175.3 g of NaCl and 88.2 g of Tri-sodium citrate in 800 ml of H_2O. Adjust pH to 7.0 with a few drops of conc. HCl. Adjust volume to 1 liter.

20xSSPE	Dissolve 210 g of NaCl, 27.6 g of $NaH_2PO_4.H_2O$, and 7.4 g of EDTA in 800 ml of H_2O. Adjust pH to 7.4 with NaOH. Adjust volume to 1 liter.
10 mg/ml EtBr	Add 1 g of ethidium bromide to 100 ml of H_2O. Stir on a magnetic stirrer of several hours to ensure that the dye has dissolved. Wrap the container in aluminum foil or transfer to a dark bottle and store at 4°C.
100% TCA solution	To a bottle containing 500 g of TCA, add 227 ml of H_2O. The resulting solution will contain 100% (w/v) TCA.
10 N NaOH	200 g/500 ml in H_2O. Store in plastic bottle.
10% SDS	Dissolve 100 g of electrophoresis-grade SDS in 900 ml of H_2O. Heat to 68°C to assist dissolution. Adjust pH to 7.2 by adding a few drops of concentrated HCl. Adjust volume to 1 liter.
Homogenization buffer	100 mM Tris-HCl (pH 8.0), 100 mM EDTA, 0.2 M NaCl.
1 M Imidazole (pH 6.8)	1.7025 g/25 ml in H_2O adjust pH to 6.8.
TE	10 mM Tris-HCl, 1 mM EDTA (pH 8.0)
TM	100 mM Tris-HCl (pH 8.0) 50 mM $MgCl_2$
Low TE	5 mM Tris-HCl (pH 8.0), 0.1 mM EDTA (pH 8.0)
1 M Tris-acetate (pH 7.8)	(M.wt. 121.14) 12.114 g/100 ml in H_2O, adjust pH to 7.8 using acetic acid.
1 M KOAC	(M.wt. 98.1) 9.81 g/100 ml in H_2O, sterilize by filtration.
1 M $Mg(OAC)_2$	(M.wt. 214.5) 21.45 g/100 ml in H_2O, sterilize by filtration.

Solution (Antibiotic)	Preparation / Composition
Ampicillin	Stock solution. 25 mg/ml of the sodium salt of ampicillin in 50% ethanol (v/v). Store in aliquots at -20°C. Working concentration. 35-50 ug/ml.
Chloramphenicol	Stock solutions. 34 mg/ml in 100% ethanol. Store at -20°C. Working concentration. For amplification of plasmids, 170 ug/ml; for selection of resistant bacteria, 30 ug/ml.
Tetracycline	Stock solution. 12.5 mg/ml tetracycline hydro chloride in 50% ethanol (v/v). Store at -20°C. Working concentration. 12.5-15.0 ug/ml. Magnesium ions are antagonists of Tetracycline. Use media without magnesium salts (e.g., LB) for selection of bacteria resistant to Tetracycline. Because tetracycline is light-sensitive, solutions and plates containing the antibiotic should be stored in the dark.

Solution (M 13)	Preparation / Composition
100 mM IPTG (isopropyl-ß-D-thiogalactopyranoside)	23.8 mg IPTG/ml in H_2O (M.Wt. 238.3) Store at -20°C.
X-gal (5-bromo-4-chloro-3-indolyl-ß-D-galactoside) or Bluo-gal	2% solution, 25 mg in 1.25 ml DMSO Store at 4°C in a dark bottle.
20% PEG/2.5 M NaCl	20 g polyethylene glycol (PEG-6000), 10 g NaCl dissovle in 100 ml distilled H_2O. Store at room temperature.
1 M K-MES (pH 6.3)	(M.wt. 195.2), 9.76 g/50 ml in H_2O using 10 M KOH to adjust pH to 6.3, filter, store at -20°C.
10 M KOH	(M.wt. 56.11) 56.11 g/100 ml in H_2O. Store in plastic bottle.
1 M RbCl	(M.wt. 120.9) 3.025 g/25 ml in H_2O, filter, store at -4°C.

0.25 M $HACoCl_3$	(M.wt. 267.46) 6.687 g/100 ml in H_2O, filter, store at -4°C.
TFB (Transformation buffer)	10 mM K-MES (pH 6.3), 100 mM RbCl, 45 mM $MnCl_2.4H_2O$, 10 mM $CaCl_2.2H_2O$, 3 mM $HACoCl_3$, filter, store at 4°C.
FSB (Frozen storage buffer)	Add 10% glycerol in TFB.
2.25 M DTT	(M.wt. 154.3) 347 mg/1 ml 40 mM KAC (pH 6.0), filter, store at -20°C.
ß-Mercaptoethanol	Usually obtained as a 14.4 M solution. Store in a dark bottle at 4°C.

Solution (Sequencing)	Preparation / Composition
0.1 M ATP	Dissolve 60 mg of ATP in 0.8 ml of H_2O. Adjust pH to 7.0 with 0.1 M NaOH. Adjust volume to 1.0 ml with H_2O. Dispense the solution into small aliquots and store at -70°C.
dNTP (or NTP)	Dissolve NTP or dNTP in Low TE (pH 8.0) directly in the shipping bottle at an concentrated concentration. Store at-70°C.
0.1 M ADP	(M.wt. 471.2 g) 47.12 mg/1 ml in H_2O.

dNTP' mixes	dATP' (ul)	dCTP' (ul)	dGTP' (ul)	dTTP' (ul)	final concentration (mM)
0.5 mM dCTP	300	15	300	300	0.0122
0.5 mM dGTP	300	300	15	300	0.0122
0.5 mM dTTP	300	300	300	15	0.0122

Chase solution	M.wt.	volume (ul)	final concentration (mM)
20 mM dATP	583	25	0.5
20 mM dCTP	569	25	0.5
20 mM dGTP	609	25	0.5
20 mM dTTP	589	25	0.5
Low TE (pH 8.0)		900	

ddNTP solution

	stock solution (mM)	working concentration (mM)
ddATP	10	0.05
ddCTP	10	0.05
ddGTP	10	0.2
ddTTP	10	0.5

This concentration is for 500 bases of sequencing gel.

40% acrylamide stock

380 g/l acrylamide and 20 g/l bisacrylamideDissolve in H_2O with stirring. Bring to final volume of 1 l. Filter through Whatmann 3 MM paper. Store at 4°C in dark bottle. Stable indefinitely.

10% (w/v) ammonium persulfate (APS)

1 g APS in 9 ml H_2O. Store at 4°C. Stable 1-2 weeks.

10 X TBE

121.1 g Tris-base, 55 g boric acid, 7.4 g Na_2EDTA. Dissolve in H_2O. Adjust to final volume of 1 l, pH should be 8.3. Store at room temperature.

Regular gel mix (500 ml)

% polyacrylamide	40% Acrylamide stock (19:1; ml)	Urea (g)	10X TBE (ml)	H_2O (ml)	Approximate mobility in bases XC	BPB
6%	75	250	50	~75	100	30
8%	100	250	50	~150	70	20
12%	150	250	50	~100	45	12
20%	250	250	50	0	25	8

Note:

1. The approximate volume of water is given to aid in dissovling the urea. Mix vigorously with a magnetic stirrer; avoid heating. Check final volume for accuracy. Filter and de-gas the solution. Store at 4°C in a dark bottle.

2. For a 8% gel and 75 ml gel mix, add 758 ul 10% APS

and 7.5 - 15 ul TEMED (10 minutes for polymerization).

3. For a 20% gel and 150 ml gel mix, add 700 ul 10% APS and 25 ul TEMED.

FDE (formamide dye mix): 0.1% xylene cyanol, 0.1% bromophenol blue, 20 mM EDTA, 90% deionized formamide. (Prepare deionized formamide with MB-1 resin and store at -20°C).

Buffer Gradient Gel
(0.5X / 2.5X, 6%)

0.5X	(1 l)
40%	150 ml
10X TBE	50 ml
urea	470 g (7.8 M)
add H_2O to	1 l

2.5X	200 ml
40%	30 ml
10X TBE	50 ml
urea	94 g (7.8 M)
40% sucrose	20 g (10%)
bromophenol blue	10 mg
add H_2O to	200 ml

	0.5X (65 ml)	2.5X (12 ml)
10% A.P.	300 ul	60 ul
TEMED	70 ul	20 ul

Note:

1. draw up in a 25 ml pipette using pipette aid: 10 ml 0.5X followed by 10 ml 2.5X and then let 5 air bubbles to pass interface to simply form gradient.

2. Running buffer: 1X TBE (prerun 30 minutes)

3. Running condition: 30x40x0.04 cm gel,
voltage: 1500-1900
mA: 30-45
watt: 60.

Solution (Hybridization)	Preparation / Composition
50X Denhardt's Solution	Ficoll 5 g polyvinylpyrrolidone 5 g BSA (Pentax Fraction V) 5 g H_2O to 500 ml Filter through a Nalgene filter. Dispense into 25 ml aliquots and store at -20°C.
Prewashing solution plaque/colony	50 mM Tris-HCl (pH 8.0), 1 M NaCl, 1 mM EDTA (pH 8.0), 0.1% SDS.
Prehybridization solution oligonucleotide probe	10X Denhardt's, 6X SSC, 0.2% SDS, 50 ug/ml yeast tRNA.
Nick-translated probe plaque/colony	50% formamide, 5X Denhardt's, 5X SSPE, 0.1% SDS, 100 ug/ml denatured, Salmon sperm DNA.
southern	6X SSC, 0.5% SDS, 5X Denhardt's, 100 ug/ml denatured, Salmon sperm DNA.
Hybridization solution oligonucleotide probe	10X Denhardt's, 6X SSC, probe
Nick-translated probe plaque/colony	50% formamide, 5X Denhardt's, 5X SSPE, 0.1% SDS, 100 ug/ml denatured, Salmon sperm DNA, denatured probe.
southern	6X SSC, 0.01 M EDTA, 5X Denhardt's, 0.5% SDS, 100 ug/ml denatured, Salmon sperm DNA, denatured probe.

Medium	Preparation / Composition
M-9 salts (10X) (minimal media)	60 g Na_2HPO_4 (M. Wt. 131.96), 30 g KH_2PO_4 (M. Wt. 136.09), 5 g NaCl, 10 g NH^4Cl, add H_2O to 1 liter.

M-9 agar (200 ml)
(minimal agar plates)

Mix 20 ml M-9 salts and 3 g of Select Agar in approximately 175 ml of water and autoclave to sterilize. Cool to approximately 55°-60°C and add 0.2 ml 1 M $MgSO^4$ (filter sterilized), 2 ml 10 mM $CaCl_2$ (filter sterilized), 2 ml 20% (w/v) glucose, and 0.2 ml 10 mg/ml thiamine (filter sterilized). Adjust volume to 200 ml with sterile water and pour into plates (about 10 plates). The plates can be stored for several months in sealed bag at 4°C.

NZCYM Medium

Per liter:

NZ amine	10 g
NaCl	5 g
yeast extract	5 g
casamino acids	1 g
$MgSO_4 . 7H_2O$	2 g

Adjust pH to 7.5 with sodium hydroxide.

Bacto-agar	15 g (for plates)
Bacto-agar	7 g (for top agar)

Maltose

Maltose (0.2%) is often added to the mediium during growth of bacteria that are to be used for plating bacteriophage lambda.

	maltose		20 g
add	H_2O	to	100 ml

Sterilize by filtration.
Add 1 ml of sterile 20% maltose solution for every 100 ml of medium.

SM
(lambda dilute solution)

This medium is used for phage storage and dilution.
Per liter:

NaCl	5.8 g
$MgSO_4 . 7H_2O$	2 g
1 M Tris.Cl (pH 7.5)	50 ml
2% gelatin	5 ml

Sterilize by autoclaving and store in 50 ml lots.

YT

Per liter:

Bacto tryptone	10 g
Yeast extract	5 g
NaCl	5 g
Bacto-agar	15 g (for plate)
Bacto-agar	7 g (for top agar)

LB (Luria-Bertani) Medium

Per liter:

Bacto-tryptone	10 g
Bacto-yeast extract	5 g
NaCl	10 g
Bacto-agar	15 g (for plate)
Bacto-agar	7 g (for top agar)

APPENDIX 2. Preparation and Composition of enzymes

Enzyme	Preparation / Composition
Pronase	20 mg/ml in H_2O. Store at -20°C. Preparations of pronase are often contaminated with DNase and RNase. These activities can be eliminated by incubating the stock solution for 2 hours at 37°C.
Proteinase K	20 mg/ml in H_2O. Store at -20°C.
Lysozyme	Stock solution. 50 mg/ml in H_2O. Dispense into aliquots and store at -20°C. Discard each aliquot after use; do not refreeze.
RNase A (Free of DNase)	Dissolve pancreatic RNase (RNase A) at a concentration of 10 mg/ml in 10 mM Tris.Cl(pH 7.5) and 15 mM NaCl. Heat to 100°C for 15 minutes and allow to cool slowly to room temperature. Dispense into aliquots and store at -20°C.
1000 units/ml RNase T_1	Dissolve in 0.1 M NaOAC (pH 5.2), 0.001 M EDTA, heat in a boiling water bath for 5 minutes to inactivate any DNases. Store at-20°C.
DNase I (Free of RNase)	Unfortunately, many commercial preparati ons of pancreatic DNase I, even those sold as "RNase-free" are contaminated by amounts of ribonuclease that are sufficient to cause significant degradation of high-molecular-weight RNA. Two methods are available to remove the contaminating RNase activity. 1. Affinity chromatography on agarose-coupled 5'-(4-aminophenyl-phosphoryl uridine 2'(3') phosphate. 2. macaloid suspension method.

APPENDIX 3. Enzyme buffer

Buffer	Preparation / Composition
10X T_4 DNA ligation buffer	250 mM Tris-HCl (pH 7.8), 100 mM $MgCl_2$, 10 mM DTT, 4 mM ATP, Store at -20°C.
T_4 RNA ligase buffer	50 mM Tris-HCl (pH 7.8), 6mM $MgCl_2$, 20 mM DTT, 1.0 mM ATP.
10X T_4 Kinase Forward buffer	500 mM Tris-HCl (pH 8.0), 100 mM $MgCl_2$, 150 mM DTT, 3.3 uM ATP.
10X T_4 Kinase Exchange buffer	500 mM Imidazole HCl (pH 6.8), 120 mM $MgCl_2$, 100 mM DTT, 2 mM ADP, 100 uM ATP.
10X TdT buffer	250 mM Potassium Cacodylate (pH 7.2), 20 mM $CoCl_2$, 10 mM DTT.
10X Agarose gel dye	0.25% Bromophenol blue, 0.25% Xylene cyanol, 25% Ficoll (type 400) in H_2O, store at room temperature.
Klenow fragment dilution buffer	10 mM Tris-HCl (pH 8.0), cool the buffer before use.
1X Bal31 buffer	12 mM $CaCl_2$, 12 mM $MgCl_2$, 20 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), 600 mM NaCl.
5X T_4 DNA polymerase buffer	165 mM Tris-HOAC (pH 7.8), 330 mM KOAC, 50 mM $Mg(OAC)_2$, 2.5 mM DTT, 0.5 mg/ml BSA.
10X Nick translation buffer	0.5 M Tris-HCl (pH 8.0), 0.05 M $MgCl_2$, 0.1 M 2-Mercap toethanol, 0.5 mg/ml nuclease free BSA.
Nick Translation dNTP mix	3 mM dNTPs except the radioactive dNTP one.

Restriction Endonuclease Buffers (1X):

	KCl (mM)	Tris-HCl$^{7.8}$ (mM)	$MgCl_2$ (mM)	DTT (mM)	BSA (ug/ml)	NaCl (mM)
Buffer A	-	25	10	1	100	50
Buffer B	-	25	10	1	100	-
Buffer C	-	25	10	1	100	100
Buffer D	25	25	10	1	100	-
Buffer E	-	25	2	1	100	50

Note 1: To digest crude prepared DNA, use same restriction buffer except use double concentration of $MgCl_2$.

2: The buffer system is established by International Biotechnologies Inc.

3: For double or more digestion, add 1 ul 1 M NaCl to each 20 ul reaction mixture after each incubation interval.

4: Use 1 unit restriction enzyme for 1 ug linear DNA(e.g., lambda phage DNA), 2-3 units restriction enzyme for 1 ug circular DNA (e.g., pBR322 DNA).

APPENDIX 4. Oligonucleotide sequence

Name	Sequence
PFK0 (N-probe)	3' <---- GT5CT6CT6GT5GT5CG --------¦ 5'
PFK1	3' <---- GTACTTGTCGTGGTACGTCGG ----¦ 5'
PFK1.5	5' ¦---- ATGATGCTGCAGCT ----------> 3'
PFK2	3' <---- CT5TT5CT5AA5AC5CC --------¦ 5'
PFK3	3' <---- T5CT5TT5ACCCT6CT ---------¦ 5'
PFK4	3' <---- AA5TACTT5TT5ACCCT --------¦ 5'
PFK5	3' <---- TT6GT6AA5CT5CT ----------¦ 5'
PFK6	3' <---- TACTTGACCTACCG ----------¦ 5'
PFK7	5' ¦---- CTGGAGCACATCTCCCG --------> 3'
PFK1N	5' ¦---- ATGACCCATGAAGAGCACCA -----> 3'
PFK7C	3' <---- AGACCCCTTCGGTGGCAGATT ----¦ 5'
PFKGAP1	3' <---- GTCGAGACGTAGGGA ----------¦ 5'
PFKGAP2	5' ¦---- TGGACCGGGCAAGGCGGC -------> 3'
PFKGAP3	3' <---- GAATGTGGCCCCCGGA ---------¦ 5'
PFK14	3' <---- AA5AA5CA7GT5CT ----------¦ 5'
PFK31	3' <---- CT6CC7CA7TT6CT ----------¦ 5'
PFK33	3' <---- CG7TACCT5CT6TT ----------¦ 5'
PFK51	3' <---- TA7AA5CT6CT6GG ----------¦ 5'
PFK52	3' <---- CT5AA5TA7AA5TT ----------¦ 5'
PFK65	3' <---- GTTGGTCACTGACTC ----------¦ 5'
PFKMC	3' <---- CC7TACTT5CG7CG ----------¦ 5'
PFK101	3' <---- CC7GT5TACGT6GT ----------¦ 5'
PFK102	5' ¦---- A5AT7GC7GC7GGAC ----------> 3'

PFK103	5' ¦---- GC7GG7GC7GA6GC ----------> 3'
PFK104	3' <---- CG7CT5CG7CG7AT ----------¦ 5'
PFK105	5' ¦---- GA5AA6TA6AC7AC7G ---------> 3'
PFK106	3' <---- CT6TT5AT5TG7TG7CT --------¦ 5'
PFK515	5' ¦---- CA5AA5ATGAA5AC ----------> 3'
PFK516	3' <---- GT6TT6TACTT6TG ----------¦ 5'
M13 sequence primer (17 mer)	5' <---- TGACCGGCAGCAAAATG --------¦ 5'
M13mp18/20 mer (EcoRI)	5' ¦---- GAATTCGTAATCATGGTC -------> 3'
M13mp18/29 mer (HindIII)	5' ¦-- CGACGGCCAGTGCCAAGCTTTTTTTTTTT ->3'

Note: 5 represents AG redundency, 6 represents TC redundency, 7 represents ACGT redundency.

APPENDIX 5. Bacterial genotype and vector size

1. Bacterial Strain

E. coli Strain	Genotype
K803	supE$^+$, supF$^+$ r$^-$K^{m-}K or r$^-$K^{m+}K
JM105	thi, rpsL, endA, sbcB15, hspR4, Δ(lac-proAB), [F', traD36, proAB, lacI^{q}ZΔM15]
JM107	endA1, gyrA96, thi, hsdR17, supE44, relA1, λ$^-$, Δ(lac-pro AB), [F', traD36, proAB, lacI^{q}ZΔM15]
JM83	ara, Δ(lac-proAB), rpsL(=strA), Ø80, lacZΔM15
HB101	hsdR$^-$, hsdM$^-$, recA13, supE44(su2$^+$), lacZ4, leuB6, proA2, thi-1(B1$^-$), Sm^R

2. Vector

Vector type	Vector	Vector size(kb)	Reference
Plasmid	pBR322	4.3	(Bolivar, 1978)
Plasmid	pBR325	5.9	(Soberon et al., 1980)
Plasmid	pBR328	4.9	(Covarrubias et al., 1981)
Plasmid	pUC9	2.7	(Vieria & Messing, 1982)
Plasmid	pUC12	2.7	(Vieria & Messing, 1982)
Plasmid	pUC18/19	2.7	(Vieria & Messing, 1982)
Bacteriophage	M13mp 10/11	7.2	(Messing, 1983)
Bacteriophage	M13mp 18/19	7.2	(Yanisch-Perron et al., 1985)
Bacteriophage	lambda Charon4A	45.4	(Blattner et al., 1977)

APPENDIX 6. Instrumentation and Equipment

Name	Source	Specification
Microcentrifuge	Brinkmann	5414/5413
Centrifuge	Brinkmann	3200
Centrifuge	International Equipment Company	HN-S
Centrifuge	Sorvall (Du Pont)	RC2-B
Ultracentrifuge	Sorvall (Du Pont)	OTD75B
Ultracentrifuge	Beckman	L5-75/L5-50B
Rotor	Sorvall (Du Pont)	SS-34/GSA
Rotor SW50.1/SW27	Beckman	VTi50/VTi65/
Slab Gel Dryer	Bio-Rad	1125B
Digital PH meter	Beckman	3500
Electrophoresis power	Northeastern Scientific supply	0-2000v/0-200mA/ 0-4000v/0-200mA
Regulated H. V. power supply	Zenith (Heath)	0-400v/0-150mA
Rotavapor	Büchi	RE120
DNA synthesizer	Applied Biosystems	380A
Personal computer	IBM	
Minicomputer	VAX	11-750
Transluminator	Ultra-Violet Produc	C-81/C-62
Water bath cooler	Sargent	S-84890
Transluminator	Fotodyne	
Camera	Polaroid	
Surveymeter	Measurements	3

Plastic bag sealer	Audion Electro	Seal boy
Vacuum Oven	Lab-line instruments	
Gyrotory water bath shaker	New Brunswick Scientific	
Sequencing electrophoresis cell	Bio-Rad	1800
Sequencing electrophoresis cell	Homemade	BRL/S0 Compatible
Agarose gel electrophoresis cell	Homemade	15x15cm gel
Controlled Environment incubator shaker	New Brunswick Scientific	
Rotalling Shaker	Warner-Chilcott Lab.	1205D3
Shaker	New Brunswick Scientific	
Incubation room	Scientific Systems	
Photoelectric Colorimeter	Klett-summer son	800-3
Laboratory Sterilizer	Amsco (General purpose)	
Gravity Convection Incubator	GCA/Precision Scientific	
Spectrophotometer	Gilford	2000
Liquid Scintillation Counter	Beckman	LS-250
Squencing Pipet	Drummond Scientific	3-00-203
Ultra Low Freezer	Revco	
Ultra Cold Freezer	Kelvinator	100

Nitrocellulose filter paper	Schleicher & Schuell	88x88 mm SQUARES 88 mm ROUND 150 mm ROUND

X-ray film	Kodak	XRP-5 XAR-5	35x43 cm 8x10 in.
Land pack film	Polariod	Type 667/55/52/57	
Slide film	Fuji Kodak	100 Ektachrome 160	

APPENDIX 7. Experimental protocols

A. Preparation of Competent Cells and Transformation

This protocol was adapted from that used by Hanahan (Hanahan, 1983). It has been used for routine cloning works and seems reliable and simple. The source of chemicals and the preparation of reagents can be easily found in APPENDIX 1.

1. Pick a single JM 107 colony off a streaked M-9 plate and grow it overnight in 3 ml YT medium at 37°C with agitation.
2. Dilute the overnight inoculum 1:100 into 50 ml YT medium and grow it at 37°C for 2.5 hours.

   @ Start to prepare soft agar 3 ml for each tube and leave it at 42°C.
3. Transfer 25 ml of culture into a sterile Oak Ridge centrifuge tube and cool on ice for 10 minutes. (Save remaining 25 ml in Falcon tube for lawn cells).

   @ Aliquot 0.2 ml lawn cells in each tube, leave it at room temperature.
4. Centrifuge cells at 4000 rpm in SS-34 rotor for 5 minutes at 4°C.
5. Resuspend cells in 8 ml cold TFB by gently vortexing, place on ice for 10 minutes.
6. Centrifuge again (as in step 4).
7. Resuspend in 2 ml cold TFB.
8. Add 70 ul DMSO, swirl and leave on ice for 5 minutes.
9. Add 70 ul of 2.25 M DTT, swirl and leave on ice for 10 minutes.

10. Add another 70 ul of DMSO, leave on ice for 5 minutes.
11. Quickly aliquot 200 ul of these competent cells into each prechilled glass test tube (ligated DNA must be added to the bottom of each tube before this step), swirl and incubate on ice for 30 minutes.

    @ Add 10 ul of 100 mM IPTG and 50 ul of 2% X-gal in each soft agar tube.
12. Heat-shock at 42°C for 90 seconds without agitation.
13. Place on ice for 90 seconds.
14. Pour soft agar into lawn cells and vortex, then pour this mixture into the competent cells and vortex, finally pour the whole mixture on YT plate.
15. Leave plates on bench until completely solidified.
16. Invert the plates and incubate at 37°C overnight.

B. Direct Gel Electrophoresis (DIGE)

This protocol was adapted from that used by Messing (Messing, 1983). It has been used to screen M13 clones and tested.

1. Centrifuge the phage stocks for one minute.
2. Remove 15 ul from the supernatants of each tube, place in a new tube, and add 1 ul of 2% SDS and 4 ul of 10 X Agarose gel loading buffer, vortex and centrifuge briefly.
3. Electrophorese the samples in a 0.8% agarose gel at 20 volts for overnight (or 70 volts for five hours). Use M13mp28 as size marker.

4. Stain the gel with 1 ug/ml EtBr and photograph it.

C. Complementary Test (C-test)

This protocol was adapted from that used by Messing (Messing, 1983). It has been used to distinguish the complementary inserts in M13 clones.

1. Take 20 ul of the supernatant from clone A and B respectively and mix it well in an Eppendorf tube.
2. Add 1 ul of 2% SDS to lyse the phage.
3. Incubate the mixture at 65°C for 90 minutes to allow hybridization to occur.
4. Cool the mixture at room temperature for 10 minutes.
5. Add 8 ul of 10 X Agarose gel loading buffer to tube and load 25 ul of the mixture to a 0.8% gel and electrophorese at 20 volts overnight (or 70 volts for 5 hours) using M13mp18 as size marker.
6. Stain the gel with ethidium bromide and photograph it.

D. Dot Hybridization

This protocol was modified from that used by Messing (Messing, 1983) and Maniatis (Maniatis et al., 1982). It was used to screen M13 library and seems simple and reliable.

1. Centrifuge the phage stocks for one minute.
2. Spot 1 ul of supernatant on a Nitrocellulose filter (NCF) paper, let the NCF air dry.

3. Denature the NCF in denaturing solution for one minute.
4. Neutralize the NCF in neutralization solution for five minutes.
5. Rinse the NCF in 2 X SSPE.
6. Bake the NCF under vacuum at 80°C for two hours.
7. Wet the NCF in 6 X SSC for five minutes.
8. Prehybridize the NCF in prehybridization solution at 42°C for four hours.
9. Hybridize the NCF in hybridization solution containing 1 x 10^7 cpm ^{32}P-labelled probe overnight.
10. Wash the NCF twice in 2 X SSC and 0.1% SDS at room temperature for 10 minutes.
11. Wash the NCF twice in 1 X SCC and 0.1% SDS at 68°C for 1.5 hours.
12. Expose the NCF under an X-ray film at room temperature or -70°C with an intensifier onvernight.

E. Mini Screen --- Quick Isolation of Small Amounts of RF

This protocol was adapted from that used by Birnboim (Birnboim & Doly, 1979). It has been used to isolate small amounts of RF or plasmid DNAs.

1. Resuspend phage stocks.
2 Inoculate 3 ml of YT medium with 5 ul of phage stock and incubate at 37°C overnight with shaking.
3. Pour 1.5 ml of the culture into a eppendorf tube, centrifuge for one minute.

4. Drain the supernatant, leaving the pellet as dry as possible.
5. Resuspend the pellet in 100 ul of ice-cold solution I.
6. Store for five minutes at room temperature.
7. Add 200 ul of ice-cold solution II. Mix the contents by invert ing the tube rapidly three times. Do not vortex. Store the tube on ice for five minutes.
8. Add 150 ul of ice-cold solution III. Mix the content by invert-ing the tube rapidly for five times. Store on ice for five minutes.
9. Centrifuge at 4°C for 5 minutes.
10. Transfer the supernatant to a fresh tube.
11. Add 0.5 ml of phenol/chloroform. Vortex and allow to stand at room temperature for 2 minutes. Vortex again, and centrifuge at room temperature for 2 minutes. Transfer the supernatant to a fresh tube. Do not transfer the interface.
12. Add 1 ml of ethanol at room temperature. Vortex and then stand at room temperature for 2 minutes.
13. Centrifuge at room temperature for 5 minutes.
14. Remove the supernatant. Stand the tube in an inverted position on a paper towel to allow all of the fluid to drain away.
15. Wash the pellet with 1 ml of 70% ethanol and recentrifuge at room temperature for 5 minutes.
16. Remove the supernatant and dry the pellets briefly by a vacuum pump.
17. Dissolve the DNA in 50 ul of TE (pH 8.0).

F. Large-Scale Preparation of RF

This protocol was adapted from that used by Birnboim (Birnboim & Doly, 1979) and Maniatis (Maniatis et al., 1982).

1. Inoculate 15 ml VT broth with a single cloning of JM 107 cells. Grow overnight at 37°C with agitation.
2. Inoculate 500 ml YT culture with 5 ml overnight culture and add 10 ul of M13 phage solution. Grow culture for 16 hours at 37°C with agitation (250 rpm).
3. Harvest the bacterial cells by centrifugation at 4000 g for 10 minutes (5000 rpm, 5 minutes for GSA rotor in 250 ml bottles) at 4°C. Discard the supernatant.
4. Wash cell pellet with ice-cold STE.
5. Resuspend the cell pellet from a 500 ml culture in 10 ml of solution I containing 5 mg/ml lysozyme. (Add lysozyme right before use.)
6. Add 20 ml of freshly made solution II and mix the contents gently. Let stand on ice for 10 minutes.
7. Add 15 ml of ice-cold solution III and mix the contents by inverting the tube sharply several times. Let stand on ice for 10 minutes.
8. Transfer to 40 ml centrifuge tube and centrifuge in a SS-34 Rotor (or its equivalent) at 12,000 rpm for 30 minutes at 4°C. The cell DNA and bacterial debris should form a tight pellet on the bottom of the tube.
9. Transfer equal quantities of the supernatant into each of two

clean 40 ml centrifuge tubes.

10. Add 0.6 volumes of isopropanol to each tube. Mix well and let stand at room temperature for 15 minutes.
11. Recover the DNA by centrifugation in a SS-34 rotor at 12,500 rpm for 30 minutes at room temperature.
12. Discard the supertantant, wash the pellet with 70% ethanol at room temperature. Discard as much ethanol as possible, then dry the pellet briefly under vacuum.
13. Dissolve the pellets in a total volume of 10 ml of TE (pH 8.0).
14. Measure the volume of the DNA solution. For every ml add exactly 1 g of solid CsCl. Mix gently until all of the salt is dissolved.
15. Add 1 ml of a solution of ethidium bromide (10 mg/ml in H_2O) for every 10 ml of CsCl solution. Mix well.
16. Centrifuge at 10,000 rpm for 10 minutes at room temperature to remove protein.
17. Measure and adjust the density. The final density of the solution should be 1.55 g/ml.
18. Transfer the CsCl solution to a tube suitable for a Ti75 rotor (or Ti50, or Ti65).
19. Centrifuge at 64,000 rpm for 16 hours at 20°C.
20. Two bands (or 3 bands) of DNA should be visible under UV light. The upper band consists of linear bacterial DNA and nicked cir cular DNA; the lower band consists of closed circular RF DNA. (The extra band represents the single stranded template DNA).

21. Collect the lower band of DNA into a tube through a needle inserted into the side of the tube.
22. Remove the EtBr by using 1-butanol saturated with water (using clinical centrifuge and the DNA will be lower layer).
23. Dialyze the aqueous phase against several changes of double distilled water.
24. Measure OD_{260}, OD_{280}.

G. Preparation of Template DNA

This protocol was adapted from that used by Sanger (Sanger et al., 1980). It has been used to prepare template DNA for sequencing. Normally, the template DNA from 1.5 ml culture is enough for 5 sequencing uses.

1. Inoculate two tubes containing 3 mls of YT medium with an individual colony from the M-9 plate of JM 107 cells and grow with shaking at 37°C for overnight.
2. Dilute enough of the overnight culture by 1:100 ratio with YT medium so that 3 mls is available for each sample.
3. Add 5 ul of supernatant of stock phage culture to tubes containing 3 mls of YT medium.
4. Grow cells with vigorous shaking at 37°C for 5 hours.
5. Transfer 1.5 ml of the culture to eppendorf tube and centrifuge for 5 minutes.
6. Pour the supernatant to a clean eppendorf tube taking care not to carry over any cells and add 200 ul of 20% PEG solution.

7. Vortex the supernatant well and leave at room temperature for 30 minutes.
8. Centrifuge for 5 minutes and remove the PEG supernatant.
9. Respin the tube for 30 seconds and remove all residual PEG.
10. Resuspend the phage pellet in 100 ul of TE (pH 8.0).
11. Add an equal volume of buffer saturated phenol, vortex well and leave for 5 minutes. Vortex and centrifuge for 2 minutes.
12. Transfer upper aqueous layer to a new tube and add 1/10 volume of 3M NaOAC, 2.5 volume of EtOH and precipitate at -70°C for 30 minutes.
13. Centrifuge for 10 minutes at 4°C and carefully pour off the ethanol.
14. Wash the pellet with 70% ethanol, centrifuge again and pour off the ethanol carefully.
15. Vaccum dry the pellet and redissolve in 25 ul of $TE^{5.0}$ and store at -20°C.
16. Take 5 ul to run a 0.8% Agarose gel to check the purity and concentration using 1 ug single stranded M13mp18 as a size marker (If necessary, take 5 ul from each sample to run complementary test. Theoretically, purified s.s.DNA give a better resolution.)

H. Preparation of Buffer Gradient Polyacrylamide Gel

This protocol was modified from that used by Biggin (Biggin et al., 1983) and BRL company (M13 cloning and sequencing manu

al). It has been successfully used to sequence the 17 kbp rabbit muscle PFK gene.

A 0.5X / 2.5 X TBE gradient across a 6% acrylamide gel is used for running the sequencing reactions. The gradient covers the lower third of the gel and is generated in a pipet by limited mixing of two solutions.

1. Carefully clean one pair of glass gel plates with detergent.
2. Siliconize the inside of the smaller plate by wiping with glass wool and allow to dry in an oven at 60°C. Wash away the remaining siliconizing solution with water and dry the plate again. (treat only one side of one plate and the siliconized plate is good for 5 to 6 uses.)
3. Right before use, wipe clean with 95% EtOH.
4. Assemble the glass plates using side and bottom spacers and bulldog clamps.
5. For a 43 cm x 35 cm gel, put 10 ml 2.5 X TBE gel mix and 65 ml 0.5 X TBE gel mix into two separate beakers.
6. Initiate polymerization by adding first 26 ul (to the 2.5 X) and 167 ul (to the 0.5 X) of 10% Ammonium Persulfate and then adding 10 ul (to the 2.5 X) and 65 ul (to the 0.5 X) of TEMED.
7. Using a pipet aid and a 25 ml pipet first take 12 ml of 0.5 X TBE gel mix and then 8 ml of 2.5X TBE gel mix. Form a rough gradient by allowing 5 air bubbles to pass through the interface.
8. Pour the gradient down the inner edge of the gel plates held at

an angle of approximately 45°. Adding the gel mix slowly and maintain an even flow. When the pipet is empty, lower the plates to horizontal and quickly take up and add the remaining 0.5 X TBE gel mix using the same pipet to fill the glass plates.

9. Put the comb in place and clamp the top edge of plates with 3 bulldog clamps and leave the gel at least for 1 hour to polymerize.
10. If the polymerized gel is stored overnight, cover the top of gel with Saran Wrap to prevent drying out.

I. Dideoxy Sequencing and Gel Running

This protocol was modified from that used by Biggin (Biggin et al., 1983) and Ahlquist (Ahlquist, lab. manual) and Sanger (Sanger, lab. manual).

1. For each template to be sequenced, Pipetts the following onto the walls of a 1.5 ml eppendorf tube:

d.d.H_2O	2.5 ul
2.5 ng/ul Primer	1 ul
T.M.	1.5 ul
template DNA	5 ul
Total Volume	10 ul

2. Centrifuge to mix and place in a 60°C water bath for 30 minutes. Cool for 10 minutes at room temperature and centrifuge to collect liquid.
3. Label uncapped eppendorf tubes with A, C, G, T along with clone

number. (If more than one clone, arrange all A tubes in one lane of microfuge rack, all C tubes in another, etc.)

4. For each of N = A, C, G, and T, mix

dNTP mix.	0.9 ul
ddNTP	0.9 ul
S^{35}dATP(11.9uci/ul)	0.2 ul
Total Volume	2 ul

for each clone (if more than one clone, multiply the volume by the number of clones).

5. Dispense 2 ul of annealed template to each of A, C, G, T tubes and then add 2 ul of A mix to the inside edge of A reaction tube. (add C, G, T mix in the same way).

6. Dilute Klenow fragment of DNA polymerase I to 0.15 units/ul in ice-cold 10 mM Tris-Cl (pH 8.0) and add 2 ul onto the walls of each tube without delay.

7. Start reactions by centrifuging briefly to mix components at room temperature.

8. After 20 minutes, add 2 ul chase solution and let reaction go for another 15 minutes.

9. At 35 minutes, add 4 ul FDE.

10. Heat tubes in a 95°C water bath for 5 minutes and quickly chill the tubes in ice.

11. Load appropriate amount of sample using Drummond Sequencing Pipet to the prerun gel and run the gel at 45 watts until the fast dye reaches the bottom of gel.

12. Remove the top buffer and bulldog clamps.
13. Tilt the glass plates to drain away the buffer left between plates.
14. Carefully separate two glass plates using spatula, leave gel lying on the larger plate.
15. Fix gel in 10% Methanol/10% Acetic Acid for 15 minutes.
16. Remove the plate with gel from the methanol/acetic acid bath.
17. Cover the gel with two sheets of 3 MM paper and invert the glass plate to transfer the gel to the 3 MM paper.
18. Cover the gel with Saran Wrap and dry the gel at 80°C for one hour using gel dryer.
19. In the dark, place a sheet of X-ray film on top of the gel and expose at room temperature for 20 hours.
20. Develop the film for 5 minutes and fix the film for another 5 minutes.

J. Dideoxy Sequencing with Microtiter Plate

This protocol was adapted from that used by Deininger (Deininger, lab. manual). This protocol was developed for large-scale sequencing works.

1. Label the first codon with clone number and the first row with A, C, G, and T respectively.
2. For each template to be sequenced, to the first line of plate, add:

	ddH_2O	2.5 ul
2.5ng/ml	Primer	1 ul
	T.M.	1.5 ul
	Template DNA	7 ul
	Total Volume -->	12 ul

3. Cover the wells with a layer of Saran Wrap and centrifuge briefly.
4. Place in an oven at 55°C for 45 minutes.
5. Let stand at room temperature for 10 minutes.
6. Centrifuge the plate briefly to concentrate any condensation and remove the Saran Wrap.
7. Distribute 2 ul of annealed mixture to the rim of each A, C, G, and T well.
8. Add 2 ul of dNTP/ddNTP^{35}SdATP mixture to the rim of each A, C, G, and T well.
9. Without delay, dilute Klenow to 0.15 units/ul with cold 10 mM Tris (pH 8.0) and add 2 ul to the rim of each A, C, G, T well.
10. Centrifuge briefly and let stay in centrifuge at room temperature for 20 minutes.
11. To each well add 0.5 mM dNTPs chase solution and centrifuge briefly to mix and leave at room temperature for 15 minutes.
12. Plate can be stored at -20°C (or -70°C) for several days, if gel will not be run immediately after this step.
13. Add to each well 2 ul of Formamide dye and centrifuge briefly.
14. Heat denature the reactions by placing the plate in an oven at

80°C for 15 minutes.

15. Cool the plate at -20°C for 2 minutes.
16. Samples are ready to load the gel at this step.

K. Restriction Endonuclease Partial Digestion

This protocol was adapted from that used by Smith (Smith & Birnstiel, 1976). It has been tested and seems reliable.

1. Use wild type lambda DNA to titrate different restriction endonuclease and find out the best condition for partial digestion.
2. Radioactively label DNA fragments at 3' or 5' end.
3. Use HindIII and EcoRI digest of wild type Lambda DNA as a size marker with 5'end labelling.
4. Use buffer A for PstI, PvuII, and HindIII; use buffer B for SstI and XmaI; use buffer C for BamHI, SalI, XbaI, and SphI; use buffer D for SmaI, buffer E for HincII.
5. Incubate 15 uls of solution (1 unit of restriction enzyme/1 ug of DNA) at 37°C for short time; for PstI: 8 minutes, for PvuII and BamHI: 40 minutes, and for all the other restriction enzymes: 30 minutes.
6. Digest 10^4 - 10^5 cpm (Crenkov counts) of each restriction endonuclease.
7. Terminate reaction by heating at 70°C for 5 minutes.
8. Add 2 uls of 10 X loading dye.
9. Load sample in a 1.2% Agarose gel (10^4 cpm for each bound will

be seen clearly after overnight exposure at room temperature).

10. Run gel electrophoresis in 1 X TBE buffer at 70 volts for 4 - 5 hours.
11. Dry the gel on a gel drier at 60°C for 30 minutes.

L. Nick Translation

This protocol was adapted from that used by Rigby (Rigby et al., 1977) and BRL company (product profile).

1. Pipette the following into a 1.5 ml microcentrifuge tube:

   5 uls of solution A (0.2 mM each of the dNTP's except dCTP)

   1 ug of DNA

   5 uls of radioactive nucleotide (alpha^{32}P-dCTP)

   make up volume to 45 uls using ddH_2O and centrifuge the mixture briefly.
2. Spot 1 ul of the solution on TLC plate for original control.
3. Add 5 uls of solution C (0.4 units/ul DNA polymerase I, 40 pg/ul Dnase I, 50 mM Tris-Cl (pH 7.5), 5mM Mg-acetate,1 mM 2-mercapto-ethanol, 0.1 mM PMSF, and 50%(V/V) glycerol, 100 ug/ml nuclease-free BSA).
4. Mix gently but thoroughly, centrifuge the mixture briefly.
5. Incubate the mixture at 15°C for 60 minutes.
6. Spot 1 ul of the mixture on TLC plate and leave the reaction tube at -20°C.
7. Run TLC in 0.75 M NaH_2PO_4 (pH 3.5) for 30 minutes.
8. Dry TLC plate and expose under an X-ray film.

9. If reaction is O.K., add 25uls of 7.5M NH_4OAC to the tube.
10. Add 150 uls of ethanol.
11. Leave the tube at -70°C for 30 minutes.
12. Centrifuge for 15 minutes.
13. Redissolve in 100 uls of double distilled water.
14. Spot 1 ul of it on a piece of 3 MM paper and count the activity using liquid scentillation counter.
15. Normally 10^7 - 10^8 cpm per ug DNA can be obtained.

M. 5' End Kinase-Forward Reaction

This protocol was adapted from that used by Chaconas (Chaconas & van de Sande, 1980).

1. Pipette the following into a 1.5 ml microcentrifuge tube:

   2.5 uls of 10 X Forward-Reaction buffer: (500 mM Tris-HCl (pH 8.0), 100 mM $MgCl_2$, 150 mM DTT, 3.3 uM ATP)

   10 uci of [gamma-^{32}P]-ATP

   5 p mole ends dephosphorylated DNA fragments.

   make up volume to 25 uls with double distilled water.
2. Spot 0.5 ul of solution on TLC plate.
3. Add 5 units of T_4 Kinase.
4. Incubate the mixture at 37°C for 30 minutes.
5. Spot 0.5 ul of the mixture on TLC plate and leave the reaction tube at -20°C.
6. Run TLC in 0.75 M NaH_2PO_4 (pH 3.5) for 30 minutes.
7. Dry TLC plate and expose under an X-ray film.

8. If reaction is O.K., add 25 uls of d.d.H_2O, 25uls of 7.5 M NH_4OAC and 150 uls of Ethanol.
9. Mix the mixture well and leave it at -70°C for 30 minutes.
10. Centrifuge for 15 minutes and resuspend it in 100 uls of d.d. H_2O.
11. Spot 1 ul of the mixture on a piece of 3 MM paper and count the activity using liquid scintillation counter.
12. Normally, 10^5 - 10^6 cpm per p mole end can be obatined.

    Note: for a 15 mer Oligonucleotide; 5 ng = 1 p mole 5'end

N. **5' End Kinase-Exchange Reaction**

This protocol was adapted from that used by Chaconas (Chaconas & van de Sande, 1980).

1. Pipette the following into a 1.5 ml microcentrifuge:

   2.5 uls of 10 X Exchange Reaction buffer: (500 mM Imidazole HCl, PH up to 6.8 with HCl, 120 mM $MgCl_2$, 100 mM DTT, 2 mM ADP, 100 uM ATP)

   50 uci of [gamma-^{32}P]-ATP

   5 p mole ends of DNA fragments

   make up volume to 25 uls with d.d.H_2O
2. Spot 0.5 ul of the solution on TLC plate.
3. Add 5 units of T_4 Kinase.
4. Incubate the solution at 37°C for 30 minutes.
5. Spot 0.5 ul of the solution on TLC plate and leave the reaction tube at -20°C.

6. Run TLC in 0.75 M NaH_2PO_4 [3.5] for 30 minutes.
7. Dry TLC plate and expose under an X-ray film.
8. If reaction is O.K., add 25 uls of d.d.H_2O, 25uls of 7.5 M NH_4OAC and 150 uls of Ethanol.
9. Mix the mixture well and leave it at -70°C for 30 minutes.
10. Centrifuge for 15 minutes and resuspend it in 100 uls of double distilled water.
11. Spot 1 ul of the mixture on a piece of 3 MM paper and count the activity using liquid scintillation counter.
12. Normally, 10^4 - 10^5 cpm per p mol end can be obtained.

O. 3' End Labelling Reaction

This protocol was adapted from that used by IBI company (commercial catalog).

1. Pipette the following into a 1.5 ml microcentrifuge tube:

   5 uls of 10 X TdT buffer: (250 mM Potassium Cacodylate (pH 7.2), 20 mM $CoCl_2$, 10 mM DTT)

   5 uci of [alpha-^{32}P] dATP

   2 p mole ends of DNA fragments

   make up volume to 50 uls with d.d.H_2O.
2. Spot 1 ul of the solution on TLC plate.
3. Add 15 units of TdT (Terminal transferase).
4. Incubate the mixture at 37°C for 30 minutes.
5. Spot 0.5 ul of the mixture on TLC plate and leave the reaction tube at -20°C.

6. Run TLC in 0.75 M NaH_2PO_4 (pH 3.5) for 30 minutes.
7. Dry TLC plate and expose under an X-ray film.
8. If reaction is O.K., add 25 uls of d.d.H_2O, 25uls of 7.5 M NH_4OAC and 150 uls of Ethanol.
9. Mix well and leave the mixture at -70°C for 30 minutes.
10. Centrifuge for 15 minutes and resuspend in 100 ul double distilled water.
11. Spot 1 ul of the mixture on a piece of 3 MM paper and count the activity using liquid scintillation counter.
12. Normally, 10^5 - 10^6 cpm per p mole end can be obatined.

P. Purification of Oligonucleotide

a. Manual De-protect Procedure of Oligonucleotide

1. Rinse the vial with 10-12 ml Acetonitrile (HPLC grade) using a 10 ml syringe.
2. Dry with Argon for 2 minutes.
3. Add 0.2-0.3 ml of Thiophenol. Allow to stand 45 minutes and remove the thiophenol.
4. Rinse with 10-12 ml Methanol (HPLC grade).
5. Rinse with 10-12 ml Acetonitrile.
6. Dry with Argon for 2 minutes.
7. Add 1 ml of Ammonium Hydroxide using a 1 ml syringe. Allow to stand 6 hours to overnight at room temperature.
8. Transfer Ammonium Hydroxide to a 1 dram vial.
9. Add the second 1 ml of Ammonium Hydroxide to the column.

Allow to stand 10 minutes to 1 hour at room temperature.

10. Add the second ml of Ammonium Hydroxide to the 1 dram vial.
11. Incubate at 55°C for 6 hours to overnight. (Note: Oligonucleotides stored in Ammonium hydroxide are stable at 4°C for months. Oligonucleotides that have been evaporated out of AmOH and resuspended in water or buffer are stable at -20°C for 1 year.
12. Evaporate to dryness using plan-shape flask and rotary evaporator.
13. Add 2 mls of 25 mM TEAB (pH 7.6) (Triethylammonium bicarbonate) to redissolve at this point and check the pH which should be above 6.

b. Desalting and Isolation of Oligonucleotide

1. Pretreat Sep-pak (a small loosely packed u-C_{18} reverse phase cartridge) as following:
   - Flush the cartridge with 10 mls of HPLC grade Acetonitrile.
   - Flush the cartridge with 10 mls of 30% Acetonitrile/25 mM TEAB.
   - Flush the cartridge with 10 mls of 25 mM TEAB.
2. Flush onto cartridge the oligonucleotide in 25 mM TEAB, maximum capacity ca. 100 A_{260} O.D. units.
3. Wash with 25 mM TEAB (5 mls for each time) until $A_{260} << 0.1$ O.D..

4. Wash with 10% CH_3CN/25 mM TEAB (ca. 5 ml).
5. Wash with 30% CH_3CN/25 mMTEAB (ca. 5 ml).
6. Evaporate collection from step 5 to dryness.
7. Add 0.5 ml of 80% Acetic Acid to remove the DMTr group for 25 minutes (orange appear).
8. Evaporate to dryness and azeotroped twice with water. (This process separates the failed sequences from the DMTr containing sequences.)
9. Dissolve in 0.5 ml of water and transfer to a 1.5 ml microcentrifuge tube and then lyphilized to dryness.
10. Wash pear-shape flask with 80 uls of formamide and transfer to microcentrifuge tube.
11. Add 10 uls of loading dye and 10 uls of NaOH.
12. Prepare a 20% polyacrylamide gel (20 cm x 40 cm x 1.3 mm) as following:

Acrylamide	30 g
Bis-acrylamide	1.5 g
Urea	72 g
10X TBE	15 ml
add H_2O to	150 ml
10% APS	700 ul
TEMED	25 ul

13. Run gel in 1X TBE, 750-1000 volts until the fast moving dye reach the bottom of gel.
14. Cover the gel with Saran Wrap and cut above and below the

tracking dyes and then remove the Saran Wrap.

15. Transfer gel onto a plastic wrap covered fluorescent TLC plate (20 x 20 cm). The band(s) can be visualized by illumination with short wave U.V. light.
16. Photograph with 2 short wave U.V. lights on the top of gel by the condition 1/4.5 x B x 14 seconds or 1/4.5 x 1 x 3 seconds.
17. Excise the slowest moving, most intense band and transfer to a 15 ml Corning plastic tube.
18. Crush to a fine powder and add 6 mls of 1 M TEAB.
19. Incubate with gentle, thorough agitation at 37°C overnight.
20. Recover supernatant by centrifugation or filtration through a sinterglass filter. (Pretreat filter using conc.HCl, wash with double dHistilled water, then rinse with 1 M TEAB, and then wash several times with double distilled H_2O).
21. Desalt via another Sep-pak and repeat step 1 to 5.
22. Evaporate to dryness and azeotroped once with water.
23. Redissolve in 0.5 ml of water and dilute 5 uls of the solution to 1 ml to read OD_{260} or scan from OD_{320} to OD_{220}.

    Note: Crude oligonucleotide without further purification via polyacrylamide is good for probe and sequencing primer. However, further purified one will give better and more clear result and pattern.

Q. Hybridization using Oligonucleotide Probe

1. Rewet NCF in 6 X SSC.
2. Prehybridize in 10 X Denhardt's, 6 X SSC, 0.2% SDS, and 50 ug/ml yeast tRNA at 55°C for 3 hours.
3. Hybridize in 10 X Denhardt's, and 6 X SSC at (T_D - 4) overnight. (Add probe right before use.)
4. Wash in 6 X SSC at room temperature first for 10 minutes. If background is still high, then temperature can be increased gradually until the maximum temperature (T_D +2).

   Note: add 2°C for each AT pair, 4°C for each GC pair, and 2°C for each redundancy, then the summary is the dissolation temp erature (T_D).

VITA

April 7, 1956	Born, Taiwan, Republic of China
1971 - 1974	Attended Hsinchu Senior High School, Taiwan, Republic of China
1974 - 1978	Attended National Taiwan University, Taiwan, Republic of China
1982	Admitted to Louisiana State University, Baton Rouge, Louisiana
1986	Candidate for Ph. D. in Biochemistry, Louisiana State University, Baton Rouge, Louisiana

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Chan Ping Lee

Major Field: Biochemistry

Title of Dissertation: Genomic Organization and Sequence Analysis of Rabbit Muscle Phosphofructokinase Gene

Approved:

Simon H. Chang
Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

N. D. Bragman

Jane M. Clayman

Doug S. Shih

David W. Talty

Ezzat S. Younathan

Date of Examination:

November 19, 1986